Genomic and transcriptomic analysis of the interaction of roses with the black spot fungus *Diplocarpon rosae*

Von der Naturwissenschaftlichen Fakultät

der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

Doktor der Naturwissenschaften

(Dr. rer. nat.)

genehmigte Dissertation

von

Enzo Nicolai Maximilian Neu (geb. Klein), M. Sc.

2018

Referent: Prof. Dr. rer. nat. Thomas Debener

Koreferent: Prof. Dr. rer. nat. Helge Küster

Tag der Promotion: 21.08.2018

Abstract

Due to the high economical importance of roses and the severe damage caused on them by the hemibiotrophic fungus *Diplocarpon rosae*, the causing agent of the black spot disease, their interaction is one best studied on cultivated roses. However, both interaction partners are lacking a public available genome sequence and very little is known about the molecular mechanisms of their interaction.

Due to this fact, in this study next generation sequencing was applied to assemble a draft genome sequence of the *D. rosae* isolate DortE4 as well as of the resistant rose genotype 88/124-46, and larger sets of transcriptome data and other external sources were used for evidence supported genome annotation with the MAKER pipeline.

Bioinformatic and experimental analyses of genes, normally occurring as single copy genes in all eukaryotic genomes, indicate that a large portion of *the D. rosae* genome is duplicated, which might indicate a whole genome duplication. Different bioinformatic approaches were combined to predict the secretome of the fungus and to identify potential effector genes in it. As expected, a significant proportion of the predicted secretome comprises enzymes for the degradation of cell wall components. 52 of 251 effector candidates matched several bioinformatic criteria and contained a Y/F/WxC motif, which is of particular interest because so far this motif was only found in effector genes of obligate biotrophic fungi. Transcriptomic data show that the majority of the predicted effector candidates are expressed during the early stages of infection and some belong to the genes with the highest expression value, making these candidates a promising starting point for various analyses.

Additionally, the transcriptomic response of the susceptible rose variety ‘Pariser Charme’ after inoculation with the powdery mildew fungus *Podosphaera pannosa* and *D. rosae* was investigated by applying the MACE approach. The comparison of the two interaction systems revealed that besides a similar response to both pathogens, pathogen-specific reactions occur. Genes related to photosynthesis and cell wall modification are down-regulated in response to *P. pannosa*, whereas genes from the phenylpropanoid and flavonoid pathways as well as of salicylic acid-signalling pathway are specifically up-regulated in response to *D. rosae*.

Key words: plant-microbe-interaction, genomics, transcriptomics

Zusammenfassung

Sternrußtau an Rosen, ausgelöst durch den hemibiotrophen Pilz *Diplocarpon rosae*, führt zu massiven Schäden an Kulturrosen. Obwohl die Interaktion mit *D. rosae* bereits gut untersucht ist, gibt es weder für Wirt, noch für das Pathogen eine veröffentlichte Genomsequenz und nur wenig ist über die molekularen Mechanismen, welche der Interaktion zugrunde liegen, bekannt.

Aus diesem Grund wurde mittels next-generation-sequencing eine Genomsequenz des *D. rosae* Isolates DortE4 und des resistenten Rosengenotyps 88/124-46 erstellt und mit dem MAKER Softwarepaket unter Zuhilfenahme externer Daten annotiert. Die bioinformatische und experimentelle Analyse von Genen, welche normalerweise als single-copy Gene in allen eukaryotischen Genomen vorliegen, zeigte, dass große Bereiche des *D. rosae* Genoms dupliziert sind, was ein Hinweis für eine Duplikation des gesamten Genoms sein kann. Verschiedene bioinformatische Ansätze wurden kombiniert, um das Sekretom des Pilzes vorherzusagen und potentielle Effektorgene in ihm zu identifizieren. Wie zu erwarten, handelte es sich bei vielen als sekretiert vorhergesagten Proteinen um Enzyme, welche Bestandteile der pflanzlichen Zellwand degradieren. Von den 251 Effektorkandidaten erfüllten 52 verschiedene bioinformatische Kriterien, u.a. das Vorhandensein des Y/F/WxC-Motivs, welches bisher ausschließlich bei Effektoren von obligat biotrophen Pilzen gefunden wurde. Transkriptomische Daten zeigen, dass die meisten Effektorkandidaten während der frühen Stadien der Infektion exprimiert sind und dass einige außergewöhnlich starke Expressionswerte erreichen, was diese Kandidaten zu einem vielversprechenden Startpunkt für verschiedene weiterführende Untersuchungen macht.

Zusätzlich wurde die transkriptionelle Reaktion der anfälligen Rosensorte „Pariser Charme“ nach Inokulation mit *D. rosae* und dem Mehltaupilz *Podosphaera pannosa* untersucht. Neben einer sehr ähnlichen Reaktion auf beide Pilze gibt es pathogen-spezifische Reaktionen, wie z.B. die Runterregulation von Genen, welche in Zusammenhang mit der Photosynthese oder der Zellwandmodifikation stehen, als Reaktion auf *P. pannosa* und die Hochregulation von Genen aus dem Phenylpropanoid- und dem Flavonoidsyntheseweg, sowie dem Salicylsäure-vermittelten Signalweg als Reaktion auf *D. rosae.*

Schlagworte: Wirt-Parasit-Interaktion, Genomik, Transkriptomik

Content

Abbreviations

8

88 genome	Draft genome of the rose genotype 88/124-46

A

Avr-factor	avirulence factor

B

bHLH	basic-helix-loop-helix
bZIP	basic domain leucine zipper

C

COI1	coronatine-insensitive 1

E

EDS1	enhanced disease susceptibility 1
EIN	Ethylene insensitive
ERF	Ethylene-response element binding factor
ERS	Ethylene response sensor
ET	ethylene
ETI	effector-triggered immunity
ETR	Ethylene response

F

FunSecKB	Fungal Secretome Knowledgebase

G

Gb	Gigabase
GO-term	Gene Ontology term

H

HR	hypersensitive reaction

I

ICS1	isochorismate synthase
ITS	internal transcribed spacer

J

JA	jasmonic acid
JAZ	jasmonate ZIM [Zinc-finger protein expressed in Inflorescence Meristem]

L

LOX	lipoxygenase
LRR	leucine-rich repeat domain
LysM	Lysin motif

M

MACE	Massive Analysis of cDNA Ends
MAMP	microbial-associated molecular pattern
MAP	mitogen-activated protein

N

NBS	nucleotide-binding site
NGS	next generation sequencing
NPR1	non-expresser of PR genes 1
NR	non-redundant

P

PAD4	phytoalexin-deficient 4
PAMP	pathogen-associated molecular pattern
PC	Rose variety ‘Pariser Charme’
PDF1.2	plant defensin 1.2
PMR4	powdery mildew resistance gene 4
PR-proteins	pathogenesis-related proteins
PTI	PAMP-triggered immunity

R

R-genes	resistance gene
R-protein	resistance protein
RNAseq	RNA sequencing
ROS	reactive oxygen species

S

SA	salicylic acid
SAGE	Serial Analysis of Gene Expression
SCF	Skp Cullin F-box containing complex

T

TIR	Toll/interleukin-1 receptor-like domain

U

UTR	untranslated region

V

Vir-factor	virulence factor

Z

ZIM	Zinc-finger protein expressed in Inflorescence Meristem

1. General introduction

1.1. Host-pathogen interaction

Due to their sessile life styles, plants have to adjust strongly to their environment. In nature plants are constantly colonised by microorganisms and have developed various mechanisms to repel them. Unlike animals, plants lack specialised mobile defence cells, like macrophages, neutrophils or dendritic cells, rather every cell of a plant has to repel a pathogen on its own (Nürnberger *et al.*, 2004). Therefore plants have developed passive barriers like the cell wall, the cuticle, trichomes or secondary metabolites as well as an active immune response which is induced in response to pathogens. Simultaneously, plant pathogenic fungi have developed various strategies to invade a plant and interfere in the plant defences including specific structures like penetration pegs, appressoria and haustoria or proteins, e.g. plant cell-wall degrading enzymes and effectors.

1.1.1. PAMP-triggered immunity and effector-triggered immunity

Plants possess an active immune system which bases on two different recognition mechanisms called PAMP-triggered immunity (PTI) and effector-triggered immunity (ETI), which can both lead to resistance of a plant to a pathogen. This model of a two-layer defence mechanism is only true for biotrophic and hemibiotrophic pathogens because ETI-mediated resistance is very rare in plants (Wang *et al.*, 2014). However, based on these two mechanisms a model of the co-evolution of host and pathogen is established which is depicted in Figure 1.1 with the example of a bacterial pathogen. This model assumes that the first line of defence is the PTI (Figure 1.1a), where highly conserved molecular patterns on the surface of a pathogen, the so called pathogen-associated molecular pattern (PAMP) or microbial-associated molecular pattern (MAMP), are recognised by membrane bound receptor-like kinases or receptor-like proteins of the plant (Bent and Mackey, 2007; Antolin-Llovera *et al.*, 2012). Typical examples for PAMPs are the bacterial flagellin protein FLS2 (Gomez-Gomez and Boller, 2000), the bacterial elongation factor Tu (Kunze *et al.*, 2004), fungal chitin or Pep13-domain of the cell wall transglutaminase and

heptaglucosides of oomycetes (Zipfel and Felix, 2005). The majority of microbes colonising plants are recognised and repelled by this mechanism.

Through natural selection, pathogens have developed effector proteins, which have among others the function to interfere with the PTI reaction as virulence factors resulting in susceptibility of the host plant (Figure 1.1b) (Jones and Dangl, 2006). A detailed description of effector proteins is given in the next paragraph (1.1.2).

In the ETI-reaction, effectors or the changes effectors inflict on the plant cell are recognised by plant receptors (Figure 1.1c), which are encoded by resistance genes (R-genes). These resistance proteins (R-proteins) possess a nucleotide-binding site (NBS) and leucine-rich repeat domain (LRR) (DeYoung and Innes, 2006). The recognition of an effector can either be direct or indirect. During an indirect recognition there is no direct interaction between effector and R-protein, only the changes on the target protein caused by the effector are recognised (Van Der Biezen and Jones, 1998). For example, the RPS5 R-protein of *Arabidopsis thaliana* recognises the AvrPphB effector of *Pseudomonas syringae* due the cleavage of the PBS1 protein by the protease activity of the effector (Ade *et al.*, 2007). The presence of R-genes is a strong selection pressure in the evolution of pathogens and can lead to the loss or modification of effectors to restore the susceptibility by avoiding the recognition through the ETI. Depending on the function of an effector, the loss or modification can also lead to a reduction in virulence (Figure 1.1d). In addition, new effectors can be developed which interfere with the ETI and suppress the defence reaction (Jones and Dangl, 2006). This co-evolutionary relationship of effector genes in the genome of the pathogen and corresponding R-genes in the genome of the host which specifically recognise one particular effector gene is called a gene-for-gene relationship (Flor, 1971) and reflects the continuous “arms race” between pathogen and host.

PAMP-triggered immunity (PTI)

Effector-triggered immunity (ETI)

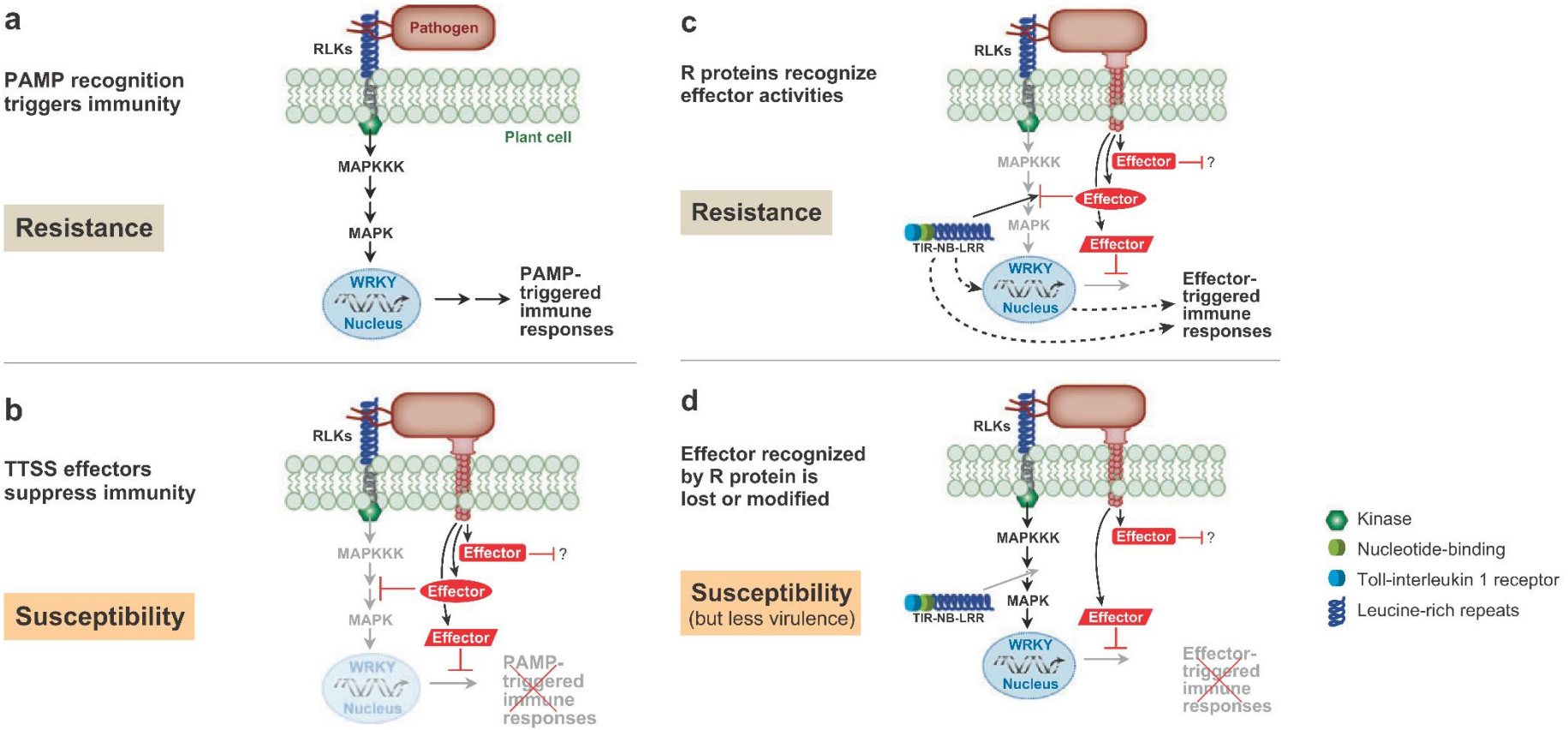


Figure 1.1: Evolutionary model of the PTI and ETI reaction. The model reflects the co-evolution of pathogen and host and is divided into four parts. (a) PAMP-triggered immunity (PTI): PAMPs on the surface of a pathogen are recognised by membrane-bound receptor-like kinases (RLKs) leading to an immune response and resistance of the host. . (b) Pathogens secret effector proteins into the host to interfere with the PTI leading to susceptibility of the host. (c) Effector-triggered immunity (ETI): The effectors are recognised by resistance proteins (R-proteins) triggering a strong immune response, often a hypersensitive response. . (d) Loss or modification of a recognised effector prevent the ETI-reaction leading again to susceptibility of the host. (Figure obtained from Bent and Mackey, 2007)

1.1.2. The role of effectors and the fungal secretome in plant pathogen interactions

As described in the previous paragraph (1.1.1), effectors are the most important factors on the pathogenic side of the pathogen-host-interaction. Due to the fact, that the group of effector proteins is very heterogeneous a clear definition is difficult. In synopsis of the literature effector proteins can be defined as proteins or toxins, which are secreted from a pathogen into the apoplast or cytoplasm of the host to acquire nutrients and modulate plant defence to promote virulence and successfully colonise the host or trigger the defence response as avirulence factors (Bent and Mackey, 2007; Catanzariti and Jones, 2010; Sperschneider *et al.*, 2015a; Lo Presti *et al.*, 2015; Selin *et al.*, 2016). By this definition it becomes clear that effectors can have a dual function as avirulence (avr) and virulence (vir) factors depending on the R-gene content in the genome of the host plant. The majority of effectors are only characterised by their avr function and the mode of action is only known for a few of them. For example the Avr2 effector of *Cladosporium fulvum* is a protease inhibitor of the Rcr3 protein of tomato which is necessary for the Cf2-mediated resistance (Luderer *et al.*, 2002; Rooney *et al.*, 2005) or the Avr4 and Ecp6 effectors of the same pathogen bind chitin to avoid its recognition and hydrolysis (van den Burg *et al.*, 2006; Jonge *et al.*, 2010). A good example for an effector functioning as a toxin is the NIP1 effector of *Rhynchosporium secalis* which induces necrotic lesions in leaf tissues of barley and other cereals (Rohe *et al.*, 1995; van't Slot *et al.*, 2007).

One of the reasons why effectors are only characterised by their avr function in most cases is that they are challenging to identify in fungal genomes (Sperschneider *et al.*, 2015a). Effectors rarely share sequence similarity due to their co-evolution with the host, but they do share some structural similarities. Effectors are often small, cysteine-rich, secreted proteins that are induced during infection (Lo Presti *et al.*, 2015; Sperschneider *et al.*, 2016). Additionally, some motifs are common in specific groups of effectors. For example, the RxLR-motif is the major identifier for effector proteins in oomycetes (Anderson *et al.*, 2015; Whisson *et al.*, 2007). Unfortunately, no such highly conserved motif has been identified in fungi, but the Y/F/WxC-motif is the only motif found in different obligate biotrophic fungi (Duplessis *et al.*, 2011; Godfrey *et al.*, 2010).

The fungal secretome is defined as the entirety of proteins that are secreted outside the plasma membrane of a cell (Girard *et al.*, 2013; Meinken *et al.*, 2014). Besides the previously described function in modulating the plant defence through effector proteins, the secretome fulfils other functions which are essential for a fungal pathogen. Due to the fact that fungi are exo-digesters, many of the secreted protein are hydrolases involved in the acquisition of nutrients. The majority of these enzymes degrade polysaccharides or proteins as a carbon and nitrogen source (Girard *et al.*, 2013). Another function, specifically important for plant pathogens, is the degradation of the plant cell wall to support the penetration process (Kubicek *et al.*, 2014).

1.1.3. Mechanisms and regulation of defence response

After the recognition of a pathogen either by PTI or ETI, signalling cascades are activated that trigger major transcriptional changes and activate various defence responses. Many of those pathway and response mechanisms are similar between a mere defence response in a compatible interaction and a resistance reaction in an incompatible interaction (Tsuda and Katagiri, 2010; Mukhtar *et al.*, 2013; Cui *et al.*, 2015; Pajerowska- Tsuda and Somssich, 2015).

The downstream signalling of the defence response contains different pathways including the production of reactive oxygen species (ROS), the activation of mitogen-activated protein kinases (MAP kinases) cascades as well as Ca^{2+} and hormone signalling (Tsuda and Katagiri, 2010; Zhang *et al.*, 2014; Torres *et al.*, 2006; Tsuda and Somssich, 2015). One of the major differences in signalling between PTI and ETI is the accumulation of ROS. PTI recognition of a pathogen triggers a rapid ROS production which is only transient, whereas the ETI reaction triggers an additional second phase of sustained ROS accumulation in a much higher magnitude (Torres *et al.*, 2006). The best studied signalling pathway is the hormone signalling pathway. Salicylic acid (SA), jasmonic acid (JA) and ethylene (ET) are the main phytohormones involved in the plant defence. It is established that SA is mainly involved in the defence response to biotrophic and hemibiotrophic pathogens, whereas JA and ET are involved in the response to necrotrophic pathogens (Figure 1.2). Both pathways are often described as antagonistic (Derksen *et al.*, 2013).

SA-signalling

Two of the major regulatory factors inducing the production of SA are enhanced disease susceptibility 1 (EDS1) and phytoalexin-deficient 4 (PAD4) which form heterodimers (Feys *et al.*, 2001; Vidhyasekaran, 2015) (Figure 1.2). It is suggested that these regulatory factors transduce ROS-derived signals to activate the synthesis of SA production mainly by the enzyme isochorismate synthase (ICS1/ SID2) (Wildermuth *et al.*, 2001). In smaller amounts SA can also be synthesised by the phenylalanine pathway (Yang *et al.*, 2015). Downstream of SA, the non-expresser of PR genes 1 (NPR1) is the key regulator of the SA mediated defence response. Under non-activated conditions NPR1 remains as an oligomeric complex in the cytosol. Increased SA-levels and changes in the cellular redox potential induce a thioredoxin-mediated release of monomers which are translocated to the nucleus where it activates further downstream response genes and recruits additional transcription factors (Boatwright and Pajerowska-Mukhtar, 2013; Pajerowska-Mukhtar *et al.*, 2013) (Figure 1.2). Thus, NPR1 is rather regulated post-translationally than transcriptionally unlike many other regulatory factors of the defence response.

JA/ET-signalling

The JA and ET pathway are generally thought to work synergistically in regulating the defence response to necrotrophic pathogens.

Free α-linolenic acid released from membranes is the basis of the synthesis of JA through different enzymes including lipoxygenase (LOX), which function as one of the marker genes of this biosynthesis pathway. The main regulator of the JA-mediated defence response is coronatine-insensitive 1 (COI1), a receptor for JA that forms the SCF^{COI1} ubiquitin ligase-complex with other proteins to degrade repressor proteins (Figure 1.2). JAZ (Jasmonate ZIM [Zinc-finger protein expressed in Inflorescence Meristem]-domain) proteins are the main repressors of JA-depending genes and a direct target of degradation through ubiquitination by the SCF^{COI1} complex. The degradation of the repressor proteins leads to the transcription of JA response genes (Vidhyasekaran, 2015) (Figure 1.2).

ET is synthesised from L-Methionine by different enzymes whereas ACC synthase is one of the most important. Without phosphorylation by MAP-kinases, this enzyme is

rapidly degraded through ubiquitin–proteasome (Liu and Zhang, 2004). The downstream regulation of the signal is mediated by different ET receptors belonging to different families: ETR (Ethylene response), EIN (Ethylene insensitive) and ERS (Ethylene response sensor) (Figure 1.2). EIN2, for instance, is not directly an ET receptor but it interacts with other ET-receptor as an essential positive regulator in ET signalling. ET leads to a phosphorylation of EIN2 that preserve it from proteasome degradation. Accumulation of EIN2 leads to further steps of the signalling pathway and in the end to the activation of ET-responsive genes (Vidhyasekaran, 2015) (Figure 1.2).

Crosstalk between hormone signalling pathways

As already mentioned, there is also crosstalk between phytohormones. For instance, mutation of the NPR1 gene leads to an enhanced expression of LOX2 and other JA marker genes, indicating that SA-accumulation suppresses the JA-ET pathway (Spoel *et al.*, 2003). Simultaneously, it was shown that JA has the potential to suppress the SA-mediated defence response (Laurie-Berry *et al.*, 2006). Nevertheless, there are also reports about a cooperative role of both phytohormone pathways in defence signalling. Additionally, other hormones like auxin, gibberellin or abscisic acid (Figure 1.2) can play a role in regulating the defence response, indicating that the signalling network is quite complex and still not fully understood (Pieterse *et al.*, 2009, Derksen *et al.*, 2013, Yang *et al.*, 2015, Verma *et al.*, 2016).

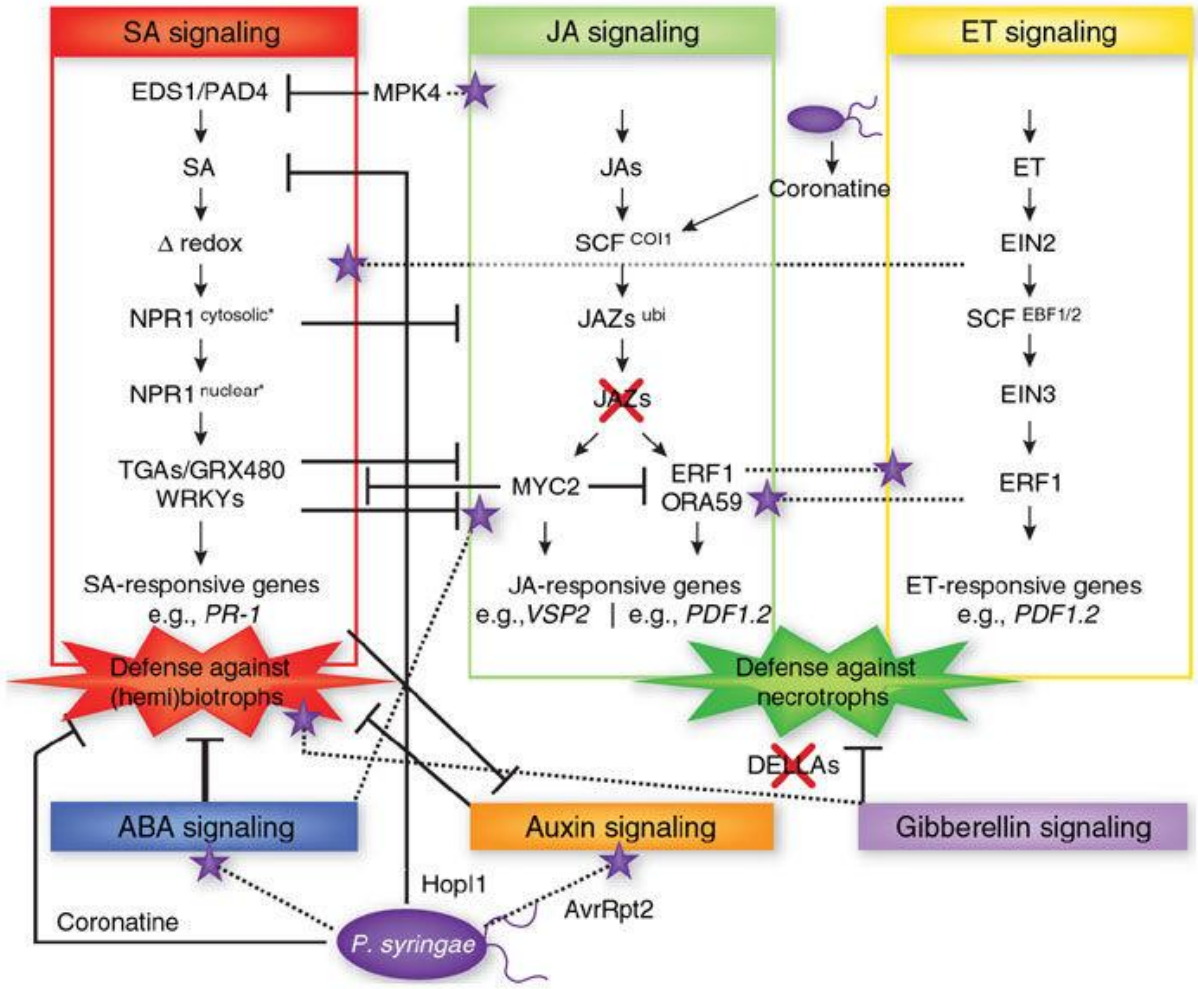


Figure 1.2: Model of the network of phytohormone signalling involved in the plant defence. The model depicts the signalling pathways of salicylic acid (SA), jasmonic acid (JA) and ethylene (ET) including some major regulatory genes as well as points of cross-talk between the pathways and other hormones. Abbreviations are as followed: disease susceptibility 1 (EDS1), phytoalexin-deficient 4 (PAD4), non-expresser of PR genes 1 (NPR1), glutaredoxin 480 (GRX480), pathogenesis-related protein 1(PR-1), Skp Cullin F-box containing complex (SCF), coronatine-insensitive 1 (COI1), jasmonate ZIM [Zinc-finger protein expressed in Inflorescence Meristem] (JAZ), ethylene-response element binding factors 1 (ERF1), octadecanoid-responsive AP2/ERF 59 (ORA59), plant defensin 1.2 (PDF1.2), vegetative storage protein 2 (VSP2), ethylene-insensitive protein 2 (EIN2), EIN3 binding F-Box proteins 1 and 2(EBF1/2), ethylene-insensitive protein 3 (EIN3), abscisic acid (ABA). (Figure according to Pieterse *et al.*, 2009)

Transcriptional regulation of defence response genes

The different signalling pathways lead to major transcriptional changes within the plant mediated by various classes of transcription factors, in particular: Ethylene-response element binding factors (ERFs), basic-helix-loop-helix (bHLH) proteins, basic domain leucine zipper (bZIP) proteins, MYB and WRKY transcription factors (Tsuda and Somssich, 2015). ERF1, for instance, is rapidly induced by ET and JA

and activates the expression of pathogen response genes (Lorenzo *et al.*, 2003). Another example is the bHLH-protein MYC2, which is one of the main regulators of JA-response genes (Lorenzo *et al.*, 2004; Kazan & Manners, 2013). TGA-bZIP transcription factors have a central role in the SA-mediated signalling pathway. Different TGAs interact with NPR1 to activate defence response genes (Zhang *et al.*, 1999; Zhou *et al.*, 2000). Besides these regulatory factors, WRKY transcription factors are involved in the regulation of various stress responses, among others the response to pathogens (Ulker and Somssich, 2004). For example, the WRKY33 is required for activating genes involved in biosynthesis of camalexin, which confers immunity to the necrotrophic fungus *Botrytis cinerea* in *Arabidopsis* (Mao *et al.*, 2011; Stefanato *et al.*, 2009). Another example is the WRKY71 transcription factor in rice which enhances the resistance to *Xanthomonas oryzae* and was shown to induce defence response genes by overexpression experiments (Liu *et al.*, 2007).

Mechanisms of the defence response

At the end of this complex signalling network stands the activation of different defence mechanisms including the up-regulation of pathogenesis-related proteins (PR-proteins), modifications of the cell-wall, induction of secondary metabolites or a hypersensitive reaction (HR) (Slusarenko *et al.*, 2000).

PR-proteins are diverse set of plant proteins which are induced in pathological or related stress situations. They are classified into 17 families based on structural or functional similarities (van Loon *et al.* 2006). The function of PR1, for instance, is not clear yet but it is one of the main marker genes for the SA-mediated defence response (Figure 1.2). Additionally, the PR2 family (endo-1,3-β-D-glucanase) and the PR5 family (thaumatin-like) are associated with this pathway. Other PR-protein families like PR4 (hevein-like protein), PR6 (proteinase inhibitor) or PR9 (peroxidase) are more associated to the JA-ET pathway. Especially the defensin PDF1.2 belonging to the PR12 family is the typical marker gene for a defence response mediated by JA and ET (Figure 1.2) (van Loon *et al.* 2006; Derksen *et al.*, 2013; Vidhyasekaran, 2015).

Another mechanism often observed as a defence response to pathogens is the modification of the cell wall, in particular the synthesis of lignin and a deposition of callose. Ton and Mauch-Mani (2004), for instance, showed in their work that callose

accumulation is one of the induced resistance mechanisms in *Arabidopsis* against necrotrophic pathogens. Another example is the powdery mildew resistance gene 4 (PMR4) of *Arabidopsis*, which codes for a callose synthase and mediates resistance to the biotrophic fungus *Golovinomyces cichoracearum* by increasing callose deposition (Ellinger *et al.*, 2013). Lignification was also observed as a defence response in different interaction systems. For instance, RNAi gene silencing experiments with wheat indicate that biosynthesis of lignin plays a critical role in the defence to the powdery mildew fungus *Blumeria graminis*. (Bhuiyan *et al.*, 2009). Similar results were also reported for the response of Chinese cabbage to *Erwinia carotovora*, which lead to increased synthesis of lignin monomers and induction of the genes of the biosynthesis pathway (Zhang *et al.*, 2007).

A third process often induced during plant defence is the induction of secondary metabolites called phytoalexins. The best studied compound is camalexin, a sulphur-containing tryptophan-derived alkaloid which is produced upon pathogen infections in Brassicaceae like *Arabidopsis* (Browne *et al.*, 1991). Many other phytoalexins are phenylalanine-derived phenylpropanoids like resveratrol of grapevine (*Vitis vinifera*) (Langcake and Pryce, 1976) or phenolic compounds produced by the flavonoid pathway like catechins which are involved in the defence response of strawberries to different pathogens (Puhl and Treutter, 2008; Yamamoto *et al.*, 2000). Additionally, terpenoids are a class of phytoalexins mainly studied in rice and maize (Piasecka *et al.*, 2015).

In addition to the described defence responses occurring in compatible and incompatible interaction, there is the HR, a form of programmed cell death which only occurs as part of the ETI-mediated resistance reaction to many pathogens. The cell death of single infected cells is meant to prevent the pathogen from spreading. ROS is of particular importance for a HR. In addition to the rapid but transient production of ROS during a PTI- and an ETI-reaction, there is a second more sustained phase of ROS production called oxidative burst associated only in the ETI-reaction. ROS has different functions during a HR: as signalling molecule, as antimicrobial compound and for cross-linking of cell wall proteins. Furthermore, the aggressive radicals cause damage to the membrane, essential proteins and the DNA, leading to a collapse of the cell and its death (Peng and Kuc 1992, Brisson *et al.*, 1994, Pontier *et al.*, 1998), (Iakimova *et al.*, 2005).

1.1.4. The defence response of members of the Rosacea family

The defence response of other species of the Rosaceae family is much better characterised compared to roses. Especially the reaction of the Rosaceae model organism strawberry (genus: *Fragaria*) was analysed in detail in different interaction systems (Amil-Ruiz *et al.*, 2011). Many of the previously described defence responses which are mainly based on analyses of model organisms like *Arabidopsis,* were also reported for Rosaceae species. Already in one of the first reports about the interaction of strawberry with *Colletotrichum fragariae* processes like cell wall thickening, pectin deposition and the accumulation of tannins were found to restrict the fungal growth (Milholland, 1982). The defence mechanisms of strawberry against pathogens have been extensively reviewed by (Amil-Ruiz *et al.*, 2011). In this review, almost all classes of PR-genes were listed to show regulation upon inoculation with pathogens in *Fragaria*. For example, expression analysis of strawberry fruits infected with *C. fragariae* showed an induction of PR5 and PR10 genes and downregulations of members of the PR2, PR9 and PR13 family (Casado-Díaz *et al.*, 2006). Several PR genes were also up-regulated in a microarray analysis of the red and white strawberry fruits infected with *C. fragariae*, among others several isoforms of major allergens belonging to the PR10-famliy (Guidarelli *et al.*, 2011). Major allergens were also amongst the highest induced genes in a very recent transcriptomic analysis of the response of *Fragaria vesca* roots to the hemibiotrophic oomycete *Pfytophtora cactorum* (Toljamo *et al.*, 2016).

In addition, both transcriptome analyses point out the induction of several genes of the flavonoid biosynthesis pathway (Guidarelli *et al.*, 2011, Toljamo *et al.*, 2016). Especially the proanthocyanidin catechin, derived from this pathway was identified as one of the major compounds of the *Fragaria* defence (Yamamoto *et al.*, 2000). However, there are also other secondary metabolites involved in the defence of strawberry. Hirai *et al.* (2000), for instance, identified triterpenes with antifungal function which are induced after the inoculation with *C. fragariae*.

Besides the response genes, some parts of the signalling network have also been analysed. It was shown that SA and JA are involved in the regulation of the defence response (Babalar *et al.*, 2007; Amil-Ruiz *et al.*, 2016) as well as ROS (Salazar *et al.*, 2007). Additionally, various transcription factors have been described to be involved in the regulation of the defence response (Amil-Ruiz *et al.*, 2011; Guidarelli *et al.*,

2011; Toljamo *et al.*, 2016). The best characterised defence related transcription factor is FaWRKY1, a homolog of the *Arabidopsis* AtWRKY75. Complementation of loss of function mutation of the *Arabidopsis* with the *Fragaria* paralog fully restored the resistance to *P. syringae* and even enhanced the resistance of wild type plants (Encinas-Villarejo *et al.*, 2009).

The described mechanisms are not limited to *Fragaria*. Other members of the Rosaceae family showed similar responses. Peach leaves (*Prunus persica*) inoculated with *Taphrina deformans* responded with ROS accumulation, callose deposition and PR-gene induction, including different PR10 genes (Svetaz *et al.*, 2017). Furthermore, different studies of apple (*Malus x domestica*) indicated an up-regulation of PR-genes as a response, in particular of genes belonging to the PR10 family (Pühringer *et al.*, 2000; Poupard *et al.*, 2003; Chevalier *et al.*, 2008; Cova *et al.*, 2017). The accumulation of flavonoids was associated with the resistance to apple-rust (Lu *et al.*, 2017). Genes belonging to the flavonoid biosynthesis pathway were also indicated to be up-regulated in a meta-analysis of transcriptomic data of resistance reactions to five different fungal pathogens (Balan *et al.*, 2018). This analysis also pointed out five different WRKYs (WRKY33, WRKY35, WRKY40, WRKY70 and WRKY75) as key factors of the defence response.

All in all, these findings indicate that many defence mechanisms can be transferred from the model systems to Rosaceae species and that some processes like the induction of PR10 genes, the secondary metabolism and some regulatory elements, are conserved within the family.

1.2. Roses as a host system

1.2.1. Taxonomy and geographic distribution of roses

As mentioned before, roses belong to the family of Rosaceae (Table 1.1), which also contains the genera of other important horticultural crops like *Malus* (apple), *Pyrus* (pear), *Prunus* (plum, peach, cherry), *Rubus* (raspberry, blackberry) and *Fragaria* (strawberry) (Potter *et al.*, 2007). The genus *Rosa* is further divided into four subgenera: *Hulthemia, Platyrhodon, Hesperrhodos* and *Rosa*, whereas the last one contains the majority of the approximate 140 to 200 known wild species. The first three subgenera comprises only two species each (Wissemann, 2003; Wissemann

and Ritz, 2005). Besides the wild species, thousands of rose cultivars are registered today as a result of to the intense breeding efforts of the last centuries. These cultivars are classified into old garden roses and modern roses. The most popular cultivar classes all belong to the modern roses with hybrid tea roses being the most popular class Zlesak, 2006).

The genus *Rosa* originates from the northern hemisphere (20-70° N. lat.) and is not indigenous to the southern hemisphere. There is evidence that roses were already cultivated in China, western Asian and norther Africa 5000 year ago (Gudin, 2003) with south-western Asia including China and Turkey as the most important centres of diversity which mainly shaped the genus as it is known today Debener and Linde, 2009). From these centres roses migrated over Europe to America. Today they are grown all over the world with large-scale cultivation of commercial cut-flower production in Chile, Zimbabwe and East Africa (Brichet, 2003).

Table 1.1: Taxonomic classification of roses. (acc. to Brands 1989)

Kingdom	**Plantae**
Subkingdom	**Viridiplantae**
Phylum	**Tracheophyta**
Subphylum	**Euphyllophytina**
Class	**Spermatopsida**
Order	**Rosales**
Family	**Rosaceae**
Genus	***Rosa***

1.2.2. The horticultural impact of roses

Roses hold great symbolic and cultural value. Due to this fact, they are entitled as the world's favourite flower (Zlesak, 2006) or as the queen of flowers Hibrand *et al.,* 2018). This is also the reason why it is the economically most important ornamental plant worldwide. In 2008, the value of the world rose production was estimated to approximately 24 billion € (Heinrichs F., 2008). More recent data of FloraHolland (2016), the world's biggest auction company for ornamentals, rank roses on the first pace of the cut-flowers, with a turnover of 746 million € and 3,377 million sold units. Simultaneously, it ranked place three of the potted plants, with 47 million units sold and a turnover of approximately 60 million €. In addition, roses are sold as garden and landscape plants. In the year 2003, 220 million garden and landscape roses

were produced worldwide (Roberts *et al.*, 2003). This value might be higher today. Besides its importance as an ornamental crop, roses are also used in smaller amounts for culinary purposes, as medical plants and as source for perfume (Gudin 2000; Zlesak, 2006, Nybom and Werlemark, 2017).

Due to its high ornamental value, ornamental characteristics like flower colour, flower shape, petal number, recurrent flowering and the overall plant habit are still the main breeding goals. Furthermore, some traits which are of particular interest depending on the use of a rose variety came into the focus of breeders very recently. For example some of the main goals in breeding cut roses are transporting quality and vase-life, which are both post-harvest characteristics. For pot roses the production of miniature cultivars with a good rooting ability became more and more important and for garden roses resistance to pathogens became one of the main breeding goals to reduce the application of pesticides (Chaanin, 2003, Debener 1999, Gudin, 2003).

1.2.3. Resistance research on roses

So far defence research on roses has mainly focused on the identification and characterisation of resistance loci of two of the most important pathogens *D. rosae,* causing, black spot and *P. pannosa*, causing powdery mildew. Thus, for black spot three resistance loci have been localised: Rdr1, Rdr2, and Rdr3 (Malek and Debener, 1998; Hattendorf *et al.*; 2004, Whitaker *et al.*, 2010). The best characterised locus is Rdr1, which contains a Toll/interleukin-1 receptor-like domain (TIR)-NBS-LRR-type resistance gene that mediates a broad spectrum resistance to various *D. rosae* isolates (Kaufmann *et al.*, 2003; Terefe-Ayana *et al.*, 2011; Menz *et al.*, 2017). For the resistance to powdery mildew three major resistance genes (Linde *et al.*, 2004; Xu *et al.*, 2005) have been identified and several quantitative trait loci regions were mapped (Dugo *et al.*, 2005, Linde *et al.*, 2006, Hosseini Moghaddam *et al.*, 2012). In addition, four homologs to the mildew resistance locus o, mediating broad-spectrum resistance through loss-of-function mutations in other species, have been mapped and characterised (Kaufmann *et al.*, 2012).

1.3. The rose pathogen *Diplocarpon rosae*

Diplocarpon rosae Wolf was first described in 1815 in Sweden by Fries (Aronescu, 1934) and is the causing agent of the black spot disease (Horst AND Cloyd, 2007), one of the most severe and damaging diseases on garden roses. The typical symptoms, which are chlorotic or necrotic spots that spread in a star shaped manor on the surface of rose leaves, gave the disease its name (Drewes-Alvarez, 2003). Beside the loss of aesthetic value leading to a reduced commercial value, fungal infection leads to a weakening of plants to the point of complete defoliation and dying of extremely susceptible genotypes. Fungal diseases on roses are mainly controlled by spraying fungicide often as a preventative measure leading to high costs, environmental contamination and the development of pesticide tolerant pathogens. Due to these reasons, breeding of resistant varieties is one of the most important goals of breeding garden roses (Debener and Byrne, 2014).

1.3.1. Taxonomic classification of the fungus

Since the description of the teleomorphic structures and its connection to the anamorphic stage *(Marssonina rosae)* of *D. rosae* by Wolf (1912), the fungus is classified into the division (Phylum) of Ascomycota. A further classification is shown in Table 1.2.

Table 1.2: Taxonomic classification of *D. rosae.* (acc. to Brands 1989)

Kingdom	**Fungi**
Subkingdom	**Dikarya**
Phylum	**Ascomycota**
Subphylum	**Pezizomycotina**
Class	**Leotiomycetes**
Order	**Helotiales**
Family	**Dermateaceae**
Genus	***Diplocarpon***
Species	***Diplocarpon rosae*** F. A. Wolf

According to Ainsworth (2008), the genus *Diplocarpon* contains six species including pathogens infecting other members of the Rosaceae family like *D. mali* (host: *Malus*), *D. mespili* (host: *Mespili*) or *D. earlianum* (host: *Fragaria*).

1.3.2. From conidia to acervuli

Propagation of the pathogen is mainly due to asexually produced conidia spreading by water and direct contact (Horst and Cloyd, 2007). The life cycle of *D. rosae* begins with the germination of these bicellular conidia within eight to nine hours after inoculation and the penetration of the cuticle (Figure 1.3). The penetration can either be direct or via the development of an appressorium. Afterwards subcuticular hyphae and intercellular hyphae are formed which penetrates epidermal and mesophyll cells, where the first haustoria are formed. The development of haustoria marks the biotrophic stage of the development of the fungus. After a phase of rapid growth, the pathogen enters its necrotrophic stage, which is characterised by the formation of intracellular hyphae that grow from cell to cell. Under favourable conditions this stage is reached after six days and is accompanied by the development of acervuli where new conidia are formed (Gachomo *et al.*, 2006).

The sexual life cycle of *D. rosae* is rarely described and it is hypothesised that it only occurs under unfavourable conditions like during winter. It is characterised by the development of ascospores via meiosis in apothecia (Frick, 1943).

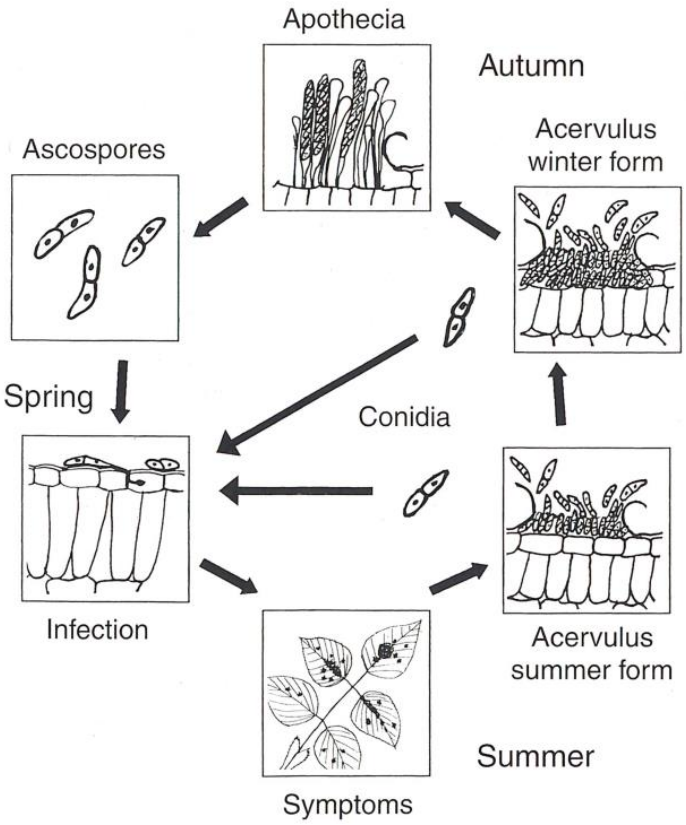


Figure 1.3: Life cycle of black spot. (Drewes-Alvarez, 2003)

1.4. Genomics and transcriptomics

1.4.1. Complexity of plant and fungal genomes and the implication for sequencing

In the last decades sequencing technologies experienced a rapid development from Sanger sequencing over second generation massive parallel sequencing to the third generation of long read single molecule sequencing, enabling scientists to sequence the genome of almost every species in an affordable way (Li *et al.*, 2017a). The first sequenced plant genome belongs to the model plant *Arabidopsis thaliana* (Arabidopsis Genome Initiative, 2000). Today more than 180 plant reference sequences have been published including model and non-model species (Jiao and Schneeberger, 2017). However, sequencing plant genomes is still a challenging task because they are in many ways more complex than microbial or mammalian genomes (Li *et al.*, 2017a; Schatz *et al.*, 2012). Plant genomes are very variable in size, ranging from small genomes like *Genlisea tuberosa* (61 Mb) and *Arabidopsis* (120 Mb) to wheat (17 Gb) or the Loblolly pine (22 Gb) (Meinke, 1998, Zimin *et al.*,

2014, Michael and VanBuren, 2015). This diversity is mainly caused by repetitive sequences like transposable elements, genome duplication, gene family enlargement and polyploidisation. Wheat, for instance, is a hexaploid crop containing approximately 90% repetitive sequences in its genome of which 60 to 80% are transposable elements (Wanjugi *et al.*, 2009).

Fungi are the most sequenced eukaryotic taxon due to their generally small genome (Ma & Fedorova, 2010). Nevertheless, their genome size can range from 8.97 Mb (*Hansenula polymorpha*) to 177.57 Mb (*Cenococcum geophilum*) (Mohanta and Bae, 2015). Reasons for this diversity are similar to those mentioned before for plant genomes. Fungal genomes can also be polyploid, contain genome duplications and enlarged gene families. Most often a large content of transposable elements is the reason for genome enlargements (Albertin and Marullo, 2012; Mohanta and Bae, 2015). The powdery mildew fungus *Blumeria graminis*, for instance, has a large genome of approximately 120 Mb which contains 64% repetitive elements (Spanu *et al.*, 2010). The genome of *Ustilago maydis* is much smaller with only 20.5 Mb and contains only 1.1% of transposable elements.

Due to the complex structure of plant and fungal genomes, generating a genome sequence is still a challenging task. The high content of repetitive and highly homologous regions often leads to highly fragmented assemblies or the collapse of similar regions. To overcome this problem, a high read coverage and sophisticated sequencing strategies are needed. The sequencing of diploid or, even better dihaploid or haploid genotypes can be used to reduce problems through the ploidy level (Michael and VanBuren, 2015). Another strategy applied in the wheat genome projects was the sequencing of a chromosome-specific BAC library (Safár *et al.*, 2010). Moreover, hybrid assemblies comprised of short read Illumina data, mate-pair data and extremely long reads derived from third generation techniques like Oxford Nanopore or PacBio single-molecule real-time (SMART) sequencing can resolve the repetitive regions. The long reads make it possible to bridge repetitive regions but still need the more accurate short read technologies for error correction (Li *et al.*, 2017a). Another way to improve genome sequences is to apply optical mapping technologies such as the nick-based method of BioNanoGenomics to generate high density physical maps for orienting and arranging of contigs Stankova *et al.*, 2016).

1.4.2. Special characteristics of the genomes of fungal pathogens

Plant pathogenic fungi contain some genomic characteristics which are highly associated with their parasitic lifestyle (Schmidt and Panstruga, 2011, Raffaele and Kamoun, 2012). Necrotrophic pathogens kill their host, while biotrophic pathogens keep their host alive by reducing damage to establish a long-term feeding relationship. This major differences lead to a specific lifestyle-associated adaption in the genome of biotrophic fungi like a reduced abundance of cell wall-degrading enzymes, non-ribosomal peptide synthetases and polyketide synthases which is more similar to symbiotic fungi (Kaemper *et al.*, 2006; Spanu *et al.*, 2010). Another feature of pathogenic fungi is the occurrence of supernumerary chromosomes which are not present in all isolates of species. A well studies example is a 1.68 Mb fragment isolated from a field isolate of *Magnaporthe oryzae* which is not present in the reference genome of the isolate 70-15. This chromosome contains candidate effector genes, some of which are avirulence factors (Yoshida *et al.*, 2009). These supernumerary chromosomal can be transferred horizontally as it was shown for *Fusarium oxysporum* (Ma *et al.*, 2010). The evolution of virulence factors and effectors are of major importance for the success of a fungal pathogen to keep the upper hand in arms race between pathogen and host. Due to this fact these factors are often embedded in very variably regions with high transposable activity or in proximity to the telomeres leading in increase in duplication, rearrangement and mutation (Cuomo *et al.*, 2007; Yoshida *et al.*, 2009; Selin *et al.*, 2016). Additionally, these factors often occur in clusters which is hypothesised to be a way of synchronising their expression (Schmidt and Panstruga, 2011, Raffaele and Kamoun, 2012).

1.4.3. The MACE technique: A novel approach for generating transcriptomic data

Today, there are several techniques available for performing large-scale expression analysis including high throughput qPCR (Microfluidic Dynamic Array), hybridization-based microarrays, massive parallel sequencing of mRNA (RNAseq) and tag-based techniques like Serial Analysis of Gene Expression (SAGE) or the Massive Analysis

of cDNA Ends (MACE) applied in this study. All with their own advantages and disadvantages.

High throughput qPCR combines the high sensitivity of qPCR with a microfluidic chip system, making the parallel analysis of several qPCR reactions possible. Nevertheless, the number of genes which can be investigated simultaneously is relatively small (max. 96, including reference genes). Its advantage is the parallel analysis of up to 96 cDNA samples, making this technique a suitable tool for the analysis of a set of candidate genes in various conditions and for the validation of other high-throughput transcriptomic data (Spurgeon *et al.*, 2008).

For many years, microarrays were the best way of generating high throughput expression data. The major advantages of microarray analyses are that these are comparatively inexpensive and that they are used for decades now, so that many biases are understood and dealt with. Nevertheless, the huge disadvantages of the technique remain. Due to the fact that immobilised probes are needed for the hybridisation, a microarray needs previous knowledge about gene models of a reference genome sequence or an EST-collection. Unpredicted genes or those which were not previously found are not included in an array, thus it cannot be detected in an expression analysis. Furthermore, there are limitations in the detection range. Expression changes of highly expressed genes might be undetectable due to the saturation of the hybridisation probes, and gene with low expression values might be undetectable because of cross-hybridisation leading to a high background. Besides that, different isoforms and splice variants are undistinguishable by this technique (Wang *et al.*, 2009; Malone and Oliver, 2011; Zheng *et al.*, 2011).

Many of the disadvantages of the microarray technique can be overcome by the RNAseq approach, in which a whole cDNA library is sequenced with next generation sequencing (NGS) techniques. Due to the fact that the complete library is sequenced, no previous knowledge is needed, making it possible to discover new genes, isoforms or splice variants. In addition, the long sequences which can be assembled are easy to annotate, because they often contain the complete coding region. However, besides these advantages, RNAseq also has some disadvantages. Due to the procedure, RNAseq data contain some biases which make a laborious normalisation necessary. For example, the amount of sequenced RNA as well as the length and the base composition of a sequenced cDNA molecule influences the

number of sequenced reads, and thus the expression value. The RNAseq technique is comparatively expensive, because of the high read coverage needed for the generation of meaningful results. Especially the detection of low expressed genes can be problematic. Up to 100 million sequenced reads can be needed for the quantification of very rare transcripts (Wang *et al.*, 2009; Zheng *et al.*, 2011; Wolf, 2013; Conesa *et al.*, 2016).

The method used in the presented study is a tag-based approach called MACE, which is an advancement of the SAGE technique. A schematic description of the MACE procedure is given in Figure 1.4. The procedure starts with the isolation of total RNA which is converted into cDNA with biotinylated oligo(dT)-primers. By means of these primers, the cDNA molecules are bound to streptavidin-coated paramagnetic beads (Kahl *et al.*, 2012). The next step is the fragmentation of cDNA molecules. This step represents the major difference between SAGE and MACE. In the SAGE procedure sequence-specific restriction enzymes are utilised to generate the fragments, whereas the fragmentation in the MACE is achieved via mechanical shearing resulting in random breakpoints (Hu & Polyak, 2006), (Kahl *et al.*, 2012). The 50 to 500 bases long fragments remain bound to the beads and unbound fragments are washed off. Starting from the breakpoint, the bound cDNA fragments are sequenced by means of NGS, generating 50–125 bp tags. These tags can be mapped to a reference sequence or used to assemble the 3’ ends of the transcripts. The major advantage of the tag-based transcriptomic approach is that every tag represents exactly one sequenced cDNA molecule, which allows a precise quantification especially of low expressed transcripts. According to the providers of the MACE technique (GenXPro), 10 to 30 times more reads would be needed to obtain a similar resolution with the RNAseq technique GenXPro, 2015). Due to this fact, MACE is also cheaper than conventional RNAseq. Nevertheless, the reduction of the cDNA molecules to their 3’ end has some drawbacks. Different isoforms or splice variants are indistinguishable if they are very similar in their 3’ end. In addition, the sequenced 3’ end often include larger parts of the untranslated region (UTR). These regions are harder to predict, thus they are often not fully included in the gene models of the reference genome, making the mapping of the reads more imprecise. In addition, if no reference sequence is available and the tags are used for a *de novo* assembly, the UTR regions interfere with the annotation because they are not part of the coding region (Wang *et al.*, 2009; Kahl *et al.*, 2012; GenXPro, 2015).

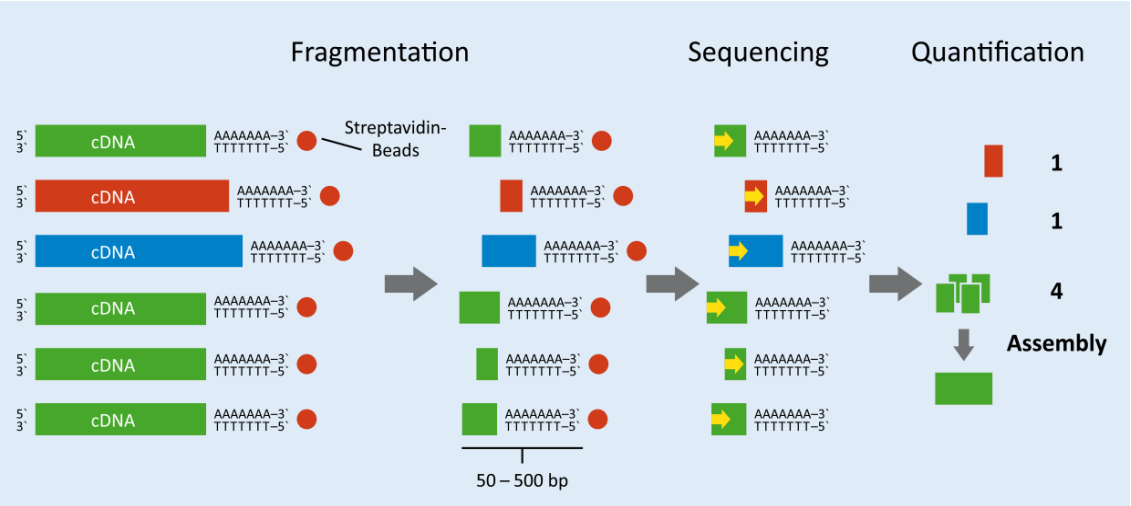


Figure 1.4: Scheme of the MACE procedure. (Kahl *et al.*, 2012)

2. Objectives of this thesis

The main goal of this thesis is to characterise the transcriptomic changes in the rose leaf during pathogen infection. To perform this analysis, the generation of basic resources in the form of a genome sequence of the fungus *D. rosae* and the rose were needed. Due to this fact, the thesis focuses on the following objectives:

I) Sequencing and characterisation of the genome of *D. rosae*
II) Prediction of the secretome and identification of effector gene candidates
III) Generation and annotation of a rose genome sequence
IV) Analysis of the rose black spot interaction by means of transcriptomic data
V) Comparison of the response to *D. rosae* and *P. pannosa*

The centre of this thesis comprises large sets of transcriptomic data, characterising a compatible rose-black spot interaction. These data were used to analyse the host response to the pathogen and to compare the results with the response to the biotrophic powdery mildew fungus *P. pannosa*. This analysis was only possible with a reliable reference like a genome sequence, which was not available at the beginning of the thesis. In addition to the sequences originated from the plant, the transcriptomic data contain sequences of the fungal pathogens which had to be separated. Due to this fact, a draft genome of the black spot fungus was generated. Additionally, the transcriptomic information was used to gain insights into the expressed proportion of the predicted fungal proteome and secretome as well as to analyse the fungal effector proteins. Hence, this thesis aims to analyse both sides of the rose-black spot interaction, which is, on the one hand, the response mechanisms of leave transcriptome to the pathogen, and on the other hand the pathogenic features of black spot fungus. These finding should be the starting point for a deeper understanding of the molecular mechanisms of the interaction between the rose and this pathogen e.g. through investigating the crosstalk between the response to biotic and abiotic stress.

3. A draft genome sequence of the rose black spot fungus *Diplocarpon rosae* reveals a high degree of genome duplication.

Enzo Neu[1], Jonathan Featherston[2], Jasper Rees[2], Thomas Debener[1]*

[1.] Institute for Plant Genetics, Leibniz University Hannover, Hannover, Germany

[2.] Agricultural Research Council, Biotechnology Platform, Onderstepoort, Pretoria South Africa

Type of authorship:	First author
Type of article:	Research article
Contribution to the article:	Planned and performed all analysis including the annotation of the genome sequence Planned and performed all experiments except for the isolation of the DNA which was performed by Thomas Debener Wrote the paper Prepared all tables and figures
Contribution of the other authors:	Jonathan Featherston and Jasper Rees performed the assembly of the sequence and wrote the material and methods part of this section Thomas Debener, Jonathan Featherston and Jasper Rees contributed in panning of the data analysis and experimental design, as well as in the revision of the manuscript
Journal:	PlosONE
Date of publication:	05.10.2017
Impact factor:	2.806
DOI:	doi.org/10.1371/journal.pone.0185310

PLOS ONE

RESEARCH ARTICLE

A draft genome sequence of the rose black spot fungus *Diplocarpon rosae* reveals a high degree of genome duplication

Enzo Neu[1], Jonathan Featherston[2], Jasper Rees[2], Thomas Debener[1]*

1 Institute for Plant Genetics, Leibniz University Hannover, Hannover, Germany, **2** Agricultural Research Council, Biotechnology Platform, Onderstepoort, Pretoria, South Africa

* debener@genetik.uni-hannover.de

OPEN ACCESS

Citation: Neu E, Featherston J, Rees J, Debener T (2017) A draft genome sequence of the rose black spot fungus *Diplocarpon rosae* reveals a high degree of genome duplication. PLoS ONE 12(10): e0185310. https://doi.org/10.1371/journal.pone.0185310

Editor: Marc Robinson-Rechavi, University of Lausanne, SWITZERLAND

Received: May 12, 2017

Accepted: September 11, 2017

Published: October 5, 2017

Data Availability Statement: The whole genome shotgun project is available in the Gene bank Database with the accession number MVNX00000000 (for further details see BioProject: PRJNA361201). The reads from the transcriptome data are available in the SRA with the accession numbers SRX2494485-SRX2494496.

Funding: This work was funded by the DFG GRK 1798 "Signaling at the Plant-Soil Interface" (GRK 1798/1) (EK, TD) and the joint research grant under the South Africa/Germany agreement on

Abstract

Background

Black spot is one of the most severe and damaging diseases of garden roses. We present the draft genome sequence of its causative agent *Diplocarpon rosae* as a working tool to generate molecular markers and to analyze functional and structural characteristics of this fungus.

Results

The isolate DortE4 was sequenced with 191x coverage of different read types which were assembled into 2457 scaffolds. By evidence supported genome annotation with the MAKER pipeline 14,004 gene models were predicted and transcriptomic data indicated that 88.5% of them are expressed during the early stages of infection. Analyses of k-mer distributions resulted in unexpectedly large genome size estimations between 72.5 and 91.4 Mb, which cannot be attributed to its repeat structure and content of transposable elements alone, factors explaining such differences in other fungal genomes. In contrast, different lines of evidences demonstrate that a huge proportion (approximately 80%) of genes are duplicated, which might indicate a whole genome duplication event. By PCR-RFLP analysis of six paralogous gene pairs of BUSCO orthologs, which are expected to be single copy genes, we could show experimentally that the duplication is not due to technical error and that not all isolates tested possess all of the paralogs.

Conclusions

The presented genome sequence is still a fragmented draft but contains almost the complete gene space. Therefore, it provides a useful working tool to study the interaction of *D. rosae* with the host and the influence of a genome duplication outside of the model yeast in the background of a phytopathogen.

cooperation in science and technology (SUA 09/030) (JF, JR, TD). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

Introduction

Fungal and oomycete pathogens are responsible for the most devastating plant diseases in temperate regions of the world [1, 2]. Many pathogens have developed sophisticated strategies to colonize and exploit their hosts by breaching various lines of defense and by manipulating the defense response of the host [3, 4]. The advent of next generation sequencing technologies has facilitated the sequencing of many fungal genomes, among which the phytopathogenic fungi have particular importance for plant diseases research. The PhytoPath [5] database contains more than 80 genomes of phytopathogenic fungi, which have been sequenced to completion, allowing crucial insights into genomic adaptations to parasitic or hemiparasitic lifestyles [6, 7]. In addition to the assembly of full genomes, draft genomes with more fragmented assemblies are very useful tools in capturing the gene space of a particular species and to identify factors relevant for host pathogen interactions [8, 9, 10].

Most sequenced phytopathogenic fungi have compact nuclear genomes of less than 50 Mb, although in exceptional cases this value can exceed 100 Mb [7].

Phytopathogenic fungi have a number of genomic characteristics that are thought to be adaptations to a parasitic lifestyle. As an example the genomes of hemibiotrophic and necrotrophic parasites encode more enzymes involved in the breakdown of complex carbohydrates than other fungal groups [11, 12]. Common to all phytopathogens are genes encoding secreted effector proteins, which are involved in the invasion process [6, 12]. As some of these effectors, or the changes they induce in the host metabolism, are recognized by the innate immune system of plants, their study is of considerable interest to both basic and applied research. Furthermore, these data can be utilized for comparative studies among diverse fungal taxa to unravel the evolutionary dynamics of adaptations involved in host-plant interactions.

Rose black spot is caused by the hemibiotrophic ascomycete *Diplocarpon rosae* (its anamorph is *Marssonina rosae*) and is one of the most damaging diseases of garden roses. Due to the world-wide trade of rose cultivars the pathogen has spread over all temperate regions of the world. The damage caused by this disease is not only indirect, due to the loss of aesthetic value of the commercial product, but also direct by weakening plants to the point of killing extremely susceptible genotypes. Propagation of the pathogen is mainly due to asexually propagated conidia spread by water and direct contact [13]. The infection starts with the germination of bicellular conidia, the formation of a germ tube and penetration of the cuticle via appressoria. During the early biotrophic stage the fungus invades the host tissue with mycelia that develop intracellular haustoria for the extraction of nutrients. This is followed by a mixed biotrophic/necrotrophic phase resulting in some tissue damage [14]. The rose black spot interaction is one of the best studied plant pathogen interactions for cultivated roses [15]. To date up to 11 pathogenic races differentiated on different sets of host plants have been characterized by various authors [15, 16] and the interaction of host and pathogen has been studied by histological and biochemical methods [14, 17, 18]. On the host side several studies addressed host resistance and a number of R-genes (resistance genes) were characterized [19, 20, 21], one of which was characterized as a TNL type resistance gene which mediates resistance against different isolates of the pathogen including DortE4 [22, 23, 24]. An interesting aspect of the pathogen biology relates to observations that indicate a low mobility of new genetic variants within and between host populations most probably due to the spread of conidia via splash water [25]. This will make disease resistance management strategies based on single R-genes interacting with single avirulence (Avr) genes more useful compared to pathosystems with extremely mobile pathogens such as powdery mildews [15].

In this study we present the draft genome sequences of the *D. rosae* isolate DortE4, which is one of the prevalent races interacting with the well-studied Rdr1 resistance locus. This

genomic sequence will serve as a working tool for the analysis of the gene space of this plant pathogen and in particular its pathogenic features. The draft can be also used as a tool to develop genetic markers for studying the population biology of the fungus. In addition to the functional information of the genome we will also analyze the genome structure and compare it to other fungi and other *D. rosae* isolates.

Results and discussion

Genome assembly

A total of 1.3 million 454 and 166.7 million Illumina paired end reads were combined to assemble the *D. rosae* genome with approximately 191-fold coverage (Table 1). These reads were assembled to contigs which were connected to 2618 scaffolds (>500 bp) with an N50 size of 243.6 kb and a total assembly length of 66.6 Mb (Table 2). These data are comparable to other fungal draft genomes [8, 9, 26] but especially the N50 value and the size of the longest scaffold (approximately 1 Mb) are outstandingly long and much larger than for others assemblies generated using similar sequence datasets. Nevertheless the genome is still very fragmented even if one takes only larger scaffolds with more than 500 bp into account.

Determination of the genome size

Flow cytometry is one of the most effective methods for the estimation of the genome size of eukaryotes and has largely replaced cytological methods. However, the adaptation of this technique to the analysis of fungal genome size is in some cases a challenging task [27, 28]. So far, we were unsuccessful to apply flow cytometry to *D. rosae* because we were not able to isolate sufficiently pure and intact nuclei. Therefore, we estimated the genome size with three different approaches based on the k-mer distribution of the reads resulting from the small insert library. An example of a k-mer distribution is depicted in Fig 1, more details are given in the supplemental file S1 File. The plot shows a clear unimodal distribution with maximum abundance at a k-mer coverage of 30.

Table 1. Sequencing statistics.

Read type	No. of reads[M]	No. of bases [Gb]	Coverage
Roche 454	1.3	0.53	6x
Illumina (short insert)	40.1	3.19	37x
Illumina (long mate pair)	126.8	12.53	147x

https://doi.org/10.1371/journal.pone.0185310.t001

Table 2. Key data of the *D. rosae* assembly.

Number of scaffolds ≥ 500 bp	**2457**
Longest scaffold	974.8 kb
Total length*	66.6 Mb
N50*	243.6 kb
L50*	84
GC-content*	47.64%
Fraction of ambiguous bases*	4.45%

*Values calculated with scaffolds larger than 500 bp.

https://doi.org/10.1371/journal.pone.0185310.t002

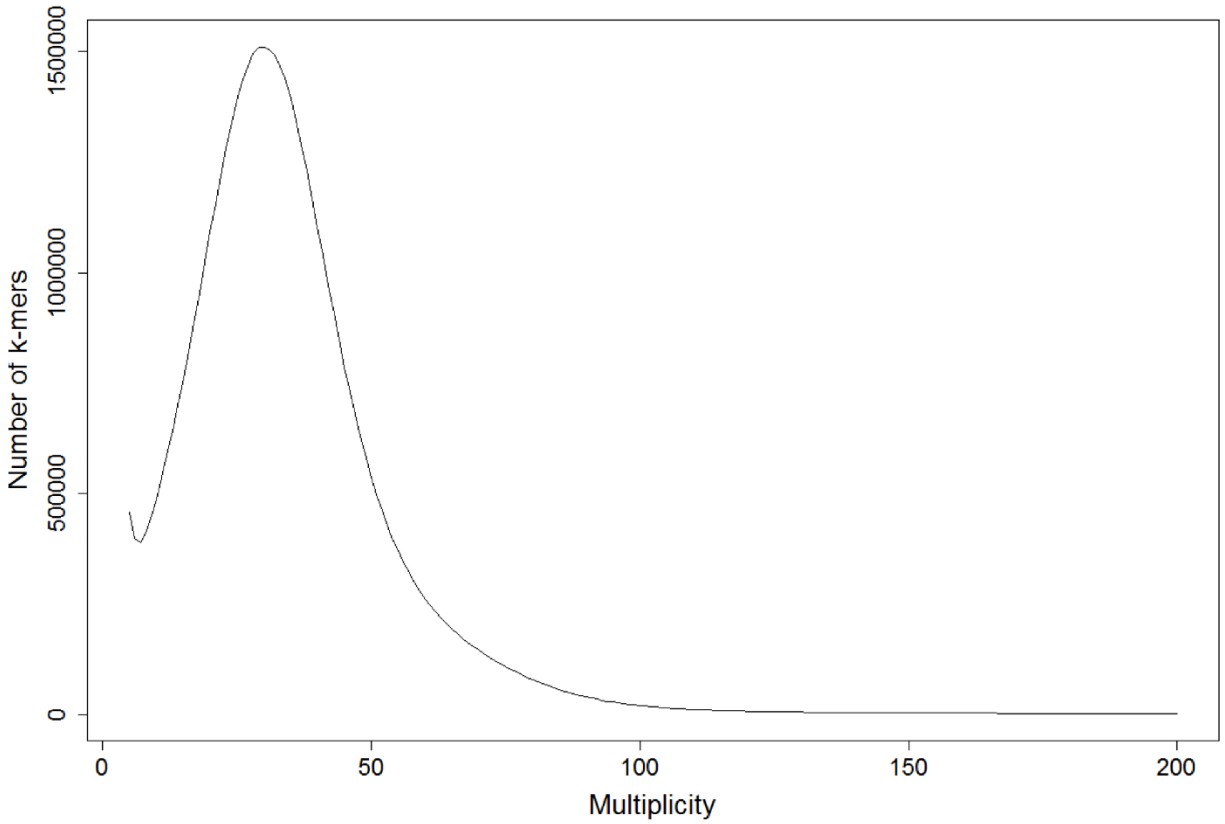


Fig 1. k-mer distribution based on the small insert Illumina library. The plot displays the number of k-mers (size 17) generated from the reads (y-axis) that occur with a given multiplicity (x-axis). The peak represents the mean k-mer depth. K-mers with extremely low frequency (<5) are not displayed and are considered to contain sequencing errors.

https://doi.org/10.1371/journal.pone.0185310.g001

The first two approaches of Li and colleagues applied for the giant panda genome [29] and Liu and colleagues for *Brassica oleracea* [30] are similar because they estimate the sequencing depth based on the k-mer coverage and calculate the genome size from this sequencing depth. Both approaches resulted in similar genome size estimations. Depending on the used k-mer size the genome size ranges from 84.99 Mb to 91.36 Mb with the approach of Li *et al.* and 83.61 to 88.6 Mb with the approach of Liu *et al.* (S1 File).

For the third approach the GenomeScope webserver [31] was used, which fits a mixture model of negative binominal model terms based on the k-mer distribution. The estimations with this software are smaller than those calculated with the other two approaches and range from 72.52 Mb to 73.53 Mb. One problem with this approach is that the software fits the model based on four peaks, which it tries to identify in the distribution. In the given data these peaks are hardly visible (S1 File), which influences the model and makes the results more erroneous.

The fungal genome size database [32] contains the data of more than 1300 ascomycetes including values derived from assembly sizes of complete genome sequencing, and from experimental methods like flow cytometry or pulse field gel electrophoresis. Taking only the experimental data into account, the mean genome size for Ascomycota is 49.4 Mb.

Unfortunately, only very few plant pathogenic fungi are included in these data. One of them is *Sclerotinia sclerotiorum* which belongs to the same order as *D. rosae*. Depending on the method used, its genome size is determined to be between 43.5 Mb [33] and 53.77 Mb [34]. Both values are much smaller than those calculated for *D. rosae*.

More data are available in the PhytoPath database [5] which collects genomic sequences and the corresponding assembly length of phytopathogens. According to this database the model organism *Magnaporthe oryzae* has an expected genome size of 41 Mb and the fully sequenced species *Botrytis cinerea* and *Marssonina brunnea*, which are closely related to *D. rosae*, with 42 Mb and 52 Mb have much smaller genomes than black spot fungus, even compared to the assembly presented here (Table 2). However the database also contains sequences of fungal pathogens which have almost the same or even larger genome sizes than *D. rosae*. The ascomycetes *Pseudocercospora fijiensis*, *Blumeria graminis* and *Verticillium longisporum* have comparable or larger genome assemblies with 73.7 Mb, 87.9 Mb and 100.5 Mb respectively. Also the *Puccinia* species, belonging to the Basidiomycota phylum, have extremely large genome sizes that range from 81.6 to 163 Mb. Despite this broad range of genome sizes, the calculated genome size of *Diplocarpon rosae* seems unexpectedly large if compared to close relatives and the majority of other phytopathogenic ascomycetes.

Annotation of the draft genome

For annotation the MAKER pipeline [35] was used with additional supporting evidence from other published closely related species and our own transcriptomic data. Table 3 summarizes the key data from the genome annotation. The number of *de novo* genes predicted ranges from 16,304 with Augustus [36] to 19,172 with the self-trained tool GeneMark-ES [37]. As a result of the MAKER pipeline, 14,004 high quality gene models were annotated on the scaffolds. The gene models have an average length of 1854 bp (S2 File), and contain an average of 3.4 exons with a mean length of 487 bp and a mean intron length of 94 bp. The predicted gene and intron length is consistent with genes found in other ascomycete genomes [38]. The other values are slightly higher than expected but as Galagan *et al.* and Mohanta and Bae showed in their reviews fungal genomes are extremely variable so that the number of genes can range from less than 6,000 to more than 27,000 and there are also examples where the mean number of exons per gene exceeds 4 [38, 39].

Automatic functional annotation of the 14,004 gene models was then performed using Blast2GO [40], which assigns a functional description to sequences based on the top 20 Blast

Table 3. Key data of the annotation pipeline.

Gene annotation with the MAKER pipeline	
Prediction tool	**Number of predicted sequences**
Augustus	16,304
GeneMark-ES	19,172
SNAP	16,931
MAKER High quality gene annotations	14,004
Functional annotation with Blast2Go	
Annotation step	**Number of annotated sequences**
Blast	13,562
Mappings (GO term assignment)	11,666
B2G: high quality annotations	10,043
InterProScan	10,839

https://doi.org/10.1371/journal.pone.0185310.t003

matches. In addition to this, the program also assigns GO-terms and performs an InterProScan [41] to provide additional characterization.

With an analysis using Blastx 96.8% of the predicted genes have a statistically significant (E-value $\leq$ 1e-10) match in the NCBI-NR database (S2 File). More than 70% of these matches originate from the fungus *M. brunnea* (S2 File) again indicating their close relationship. Most of Blast matches display similarities of more than 50% to the target sequences with an average value of 73.2% (S3 File). GO terms were assigned to 11,666 of the sequences and only 1623 of these GO annotations did not pass the internal quality control of Blast2Go. In this way more than 50,000 high quality GO annotations were assigned to 10,043 sequences. In addition to the GO-annotation, 2460 genes were assigned to an ENZYME EC number [42] and the InterProScan linked functional information for 10,839 gene models (S3 File).

Transcriptomic analysis of the predicted gene space

With a combination of the MACE (Massive Analysis of cDNA Ends) and RNA-Seq a transcriptomic dataset was generated covering the first stages of the compatible interaction between black spot and roses to analyze transcription of the predicted black spot genes. The data comprises three time points (0, 24, 72 hpi) represented by three (MACE) to six (MACE and RNA-Seq) datasets each set based on independent inoculation experiments of DortE4 with the susceptible variety "Pariser Charme". These sequence data were mapped to the annotated genomes to analyze the portion of predicted transcripts expressed over the first three days of the infection process.

A total of 12,396 (88.5%) of the predicted genes were represented by expressed transcripts, with the majority (8095) of these genes exclusively detectable three days after inoculation (Fig 2, S4 File). As our data only cover the first phase of the fungal development, including the formation of haustoria while other developmental stages of the fungus such as the formation of plectenchymatic tissues, acervuli and conidia as well as the sexual stages were not included, so it is not surprising that not all genes were found to be expressed. Another reason for undetected transcripts is the fact that in total only 1.5 million MACE and 2.4 million RNA-Seq reads matched the genome of *Diplocarpon*. The majority of the sequenced reads originated from the rose transcriptome because at early stages of infection the fungal biomass is comparable small to the amount of plant tissue. However, at later stages where the growing fungal mycelium has increased in biomass, many more transcripts could be detected. The fact that the majority of the transcripts occur three days after inoculation is most probably due to this effect and the development of additional/new organs like e.g. haustoria which are not detectable until 24 h after inoculation.

Another interesting point is that with the RNA-Seq data 12,145 transcripts were detected at 72 hpi whereas with the MACE data only 7,407 (Fig 2). This might be due to the MACE procedure where only the 3' ends, especially UTR regions of the mRNA molecules, were sequenced. But these regions are especially hard to predict by the *ab initio* gene prediction tools, indicating that the current annotation might underestimate the UTR regions of the genes. At the same time, there are transcripts that were only detected by the MACE data. These might be low abundance genes, which can be better detected by MACE than RNA-Seq [43].

Pathogenic features of the genome

To get a better insight into the putative functions of the annotated sequences in the pathogenic process, two databases (the Phi-base and the CAZy database) were used to gather additional information. The Phi-base (pathogen-host interaction database) [44, 45] contains virulence, pathogenicity and effector genes of fungi, oomycetes and bacteria. The entries are categorized

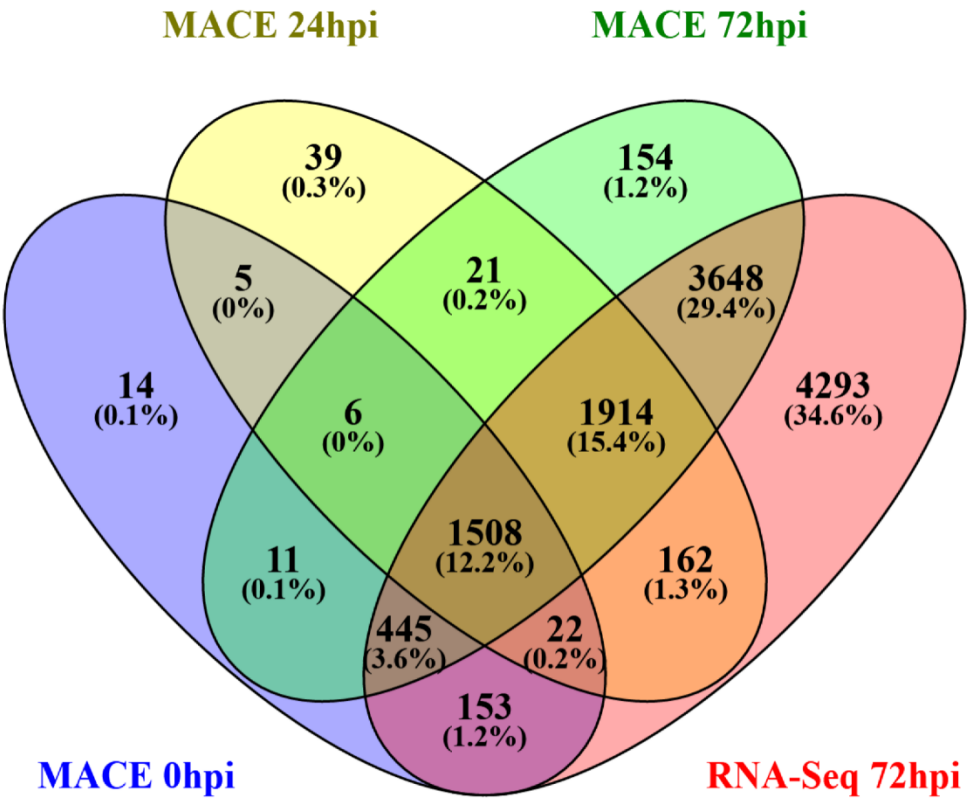


Fig 2. Venn diagram of expressed genes during early stages of the infection process (0, 24, 72 hpi). Three independent inoculation experiments were performed and genes are considered to be expressed at a time point, if they show expression in at least two of the three experiments. MACE and RNA-Seq data for 72 hpi were generated from the same samples.

https://doi.org/10.1371/journal.pone.0185310.g002

based on information about the influence of mutagenesis experiments on the virulence of the pathogens. Using Blastp (E-value $\leq$ 1e-10) a total of 3683 *D. rosae* predicted genes matched sequences in the Phi-database. Of these 17 are classified as known effectors, 346 showed a loss of pathogenicity in mutagenesis experiments and 1530 a reduction of virulence (S5 File).

58% of these sequences are detected by the MACE analysis and 90% by the RNA-Seq data, indicating that they are expressed during the early stages of the infection. Of particular interest for further analysis are the 15 expressed homologs of known effectors because they can give a closer insight into the interaction between the pathogen with its host. Finding functional effectors can give information about how the pathogen modulates the host immune response or acquires nutrients from it. They can also be useful tools for finding new R-genes and to decode their mode of action. But also the other virulence factors can give valuable information about the pathogenic strategy of a hemibiotrophic lifestyle.

The CAZy database [46] was used to detect groups of genes that code for carbohydrate degrading, modifying and synthesizing enzymes. This includes cell wall degrading enzymes (CWDEs) like cellulases, glucanases, xylanases, pectin lyases and other hydrolytic enzymes. With the dbCAN webserver [47] 724 of the predicted genes were assigned to protein families from the CAZy database (S6 File) including 285 glycoside hydrolases (GH) from 49 families,

129 carbohydrate esterases (CE), 168 glycosyl transferase (GT), 32 polysaccharide lyases (PL), 38 sequences carrying a carbohydrate-binding module (CBM) and 104 with auxiliary activity (AA). The last family contains among others lignin peroxidases (AA2), cellobiose dehydrogenases (AA3) and lytic polysaccharide monooxygenases (AA9) [48]. Kubicek and colleagues compared the CWDE content of 35 saprophytic, necrotrophic and hemibiotrophic fungi [11]. In general the content of CWDEs in the *D. rosae* genome corresponds with other hemibiotrophic fungi, but some features like the large number of cellulases, especially of the family GH5, and many pectin degrading enzymes (GH28, GH78, PL1, PL3, PL4, PL9 and PL11) match more with necrotrophic fungi. It is also noticeable that the genome contains fewer genes for hemicellulolytic enzymes than other fungi. This might be due to differences in the cell wall composition of roses or may just be another reflection of the extreme diversity of fungal genomes.

Of particular interest are the five sequences, which are assigned to the LysM-family (CMB50, IPR018392). This protein class might contain effector proteins. The Ecp6 protein of *Cladosporium fulvum* has been shown to interfere with the chitin induced defense response due to the chitin binding function of this domain [49]. Similar functions are reported for LysM-proteins from *M. brunnea*. It is shown that LysM-proteins of this fungus can interact with chitin in a way similar to Ecp6. In addition, expression of two different fungal LysM-protein with plant secretion signals in *A. thaliana* led to a reduced induction of *PDF1.2* expression in response to chitin, indicating a reduced chitin dependent defense response [50]. Therefore, the five expressed LysM-genes (S4 File) are useful starting points for further analysis and can complement the genes identified from the PHI-database.

Repeat structure of the *Diplocarpon* genome

An analysis of the repetitive fraction of the genome (Table 4) revealed an unexpectedly small fraction (15%) of repeats in the *Diplocarpon* genome in comparison to its large predicted genome size. In general this amount of repetitive and transposable elements is large compared to other fungi. The reference genome of the model fungus *M. oryzae* contains only 9.7% of repetitive elements and the assemblies of the related fungi *S. sclerotiorum* and *B. cinerea* contain only 7% respectively less than 1%. However, these genomes are all much smaller than that of *D. rosae*. In many reported cases genome enlargement is due to an expansion of the noncoding DNA especially transposable elements (TEs) and repeats [51, 52]. The genome of *Blumeria graminis* for example has a size of around 120 Mb and consists of 64% TEs [53]. The same was observed for the genome of *Tuber melanosporum* with a size of 125 Mb and a TE portion of 58% [54]. Even the smaller genome (52 Mb) of the closeted sequenced relative *M. brunnea* has a much higher repeat content of 42% [55]. There is a possibility that in the presented sequence some TEs collapsed during the assembly, which might be an explanation for the difference between the assembly length (66.6 Mb) and the predicted genome size (73–91 Mb).

A total of 7.8% of the genome assembly is composed of mobile genetic elements with LTR elements representing the largest group (7.10%) (Table 4). All LTR elements belong to the Copia or Gypsy family (S7 File) which is not surprising, because these families are the most often reported types of transposable elements in fungi [51]. The other class of retrotransposons, the non-LTR elements (SINEs and LINEs), represent only 0.24% of the assembled sequence in which no SINE elements were detected. DNA transposons and rolling cycle elements tend to play a secondary role with less than 0.5% of the genome assembly each. Those results reflect the above mentioned general trend that the amount of transposable elements is larger than that of other fungal genomes but not as large as expected. For the genome of *M. oryzae* only 3.4% of LTR retrotransposons were reported, for *S. sclerotiorum* and *B. cinerea* less

Table 4. Distribution of transposable elements and repeats in the genome sequence.

Type	Number	Length [bp]	Percent of the assembled genome sequence
SINEs	0	0	0
LINEs	163	164,700	0.24
LTR elements	8428	4,812,228	7.10
DNA transposons	562	195,740	0.29
Rolling-circles	140	114,992	0.16
Unclassified	11309	3,748,693	5.53
Simple repeats	22241	902,353	1.33
Low complexity	2545	127,064	0.19

SINE: short interspersed nuclear element
LINE: long interspersed nuclear element
LTR: long terminal repeat

https://doi.org/10.1371/journal.pone.0185310.t004

than 1% [55]. Much larger amounts of these elements were detected in the genome of *M. brunnea* and *B. graminis* with approximately 26% respectively 12.4%. For non-LTR retroposons and DNA transposons the picture is different because out of the mentioned fungi only *B. cinerea* contains smaller amount of these elements than the analyzed assembly. All other sequences contain between 1.1% and 1.9% of these elements. *B. graminis* is an exception here because with more than 20% an outstanding amount of non-LTR elements is reported in its genome [51, 53]. The trend for the unknown repetitive elements as well as the more undefined simple repeats and low complexity regions is that these elements occur in higher rates in the *D. rosae* sequence compared to other fungal genomes but not in rates as high as in *M. brunnea* and *B. graminis*.

Altogether, those findings indicate that the content of repetitive elements is higher in the presented assembly than in other fungal genomes but the large genome size cannot be explained by these elements alone.

Apart from the relevance of the repeat content for the genome structure, this information can also be used to generate simple sequence repeat (SSR) markers which can used in studies about the population biology of *Diplocarpon*. Therefore, the genome sequence of DortE4 and related plant pathogenic fungi was screened with the SSR locator tool [56]. The results, (Table 5) indicate that the *D. rosae* genome contains a total of 8242 SSR motifs, which is less than *M. brunnea*. This corresponds to the lower repeat content as mentioned before. Interestingly *M. oryzae*, which contains only 9.7% repetitive sequences in its 41 Mb sized genome, has almost the same content of SSRs as *D. rosae*. This indicates that repeat structure and SSR content do not correlate well in these organisms.

Neither is there a correlation between the genome size and the number of SSR motifs, as can be seen by comparing the genome of *M. brunnea* with that of *D. rosae*. The genome of *M.*

Table 5. SSR content of the genome.

Species	Assembly size [Mb]	Mono	Di	Tri	Tetra	Penta	Hexa	Hepta	Octa
D. rosae	66.6	6150	1629	350	58	21	16	6	12
M. brunnea	52	10618	1229	494	196	66	63	27	36
B. cinerea	41.2	3024	545	162	89	24	33	25	20
S. sclerotiorum	38.5	2403	112	88	68	36	64	53	23
M. oryzae	41	8375	277	102	12	3	19	3	5

https://doi.org/10.1371/journal.pone.0185310.t005

brunnea is smaller than that of *D. rosae*, but contains more SSRs. The same pattern emerges when comparing the genomes of *B. cinerea* and *S. sclerotiorum*, which have comparable genome sizes of 42 Mb and 38 Mb respectively but huge differences in SSR content. The same was observed in a study by Karaoglu, Lee and Meyer [57], which examined SSR motif abundance in nine different fungal genomes of very diverse sizes. They found no correlation between the genome size and SSR content.

Coverage of the gene space based on BUSCO analysis

To estimate if the genome assembly, as well as the predicted transcriptome, contains the whole gene space of *Diplocarpon*, an analysis was performed with the BUSCO pipeline [58]. This pipeline examines the presence of single-copy orthologs, which are conserved between almost all species within one phylogenetic clade. For fungi this dataset comprises 1438 genes. For the genomic sequence of *D. rosae*, 96.3% of the BUSCO orthologs were detected as complete sequences, 2.8% were fragmented and less than 1% are missing (Table 6). Almost the same results were found for the predicted transcriptome with 96.5% full length sequences and 1.5% fragmented sequences. Only 1.9% of the orthologs are missing in the predicted transcriptome, although they are detectable at the genomic level but with no gene model predicted. Therefore, the gene prediction might lead to a slight underestimation of the gene space.

Comparing these results to the BUSCO analysis of other sequenced fungal genomes indicates that our draft genome covers the gene space almost as completely as the fully sequenced genomes of *M. brunnea* [55], *B. cinerea*, *S. sclerotiorum* [59] and the model organism *M. oryzae* [60], which contain between 96.5% and 98.9% of the BUSCO orthologs as full length sequence and between 0.7% and 2.7% as fragmented sequences.

However, the genomes differ in the number of orthologs found in multiple copies that normally occur as single copies. In total, 77% of the complete BUSCO genes in the *D. rosae*

Table 6. Comparison of BUSCO results of the *D. rosae* sequence and other fungi.

	Complete BUSCOs	Complete single-copy BUSCOs	Complete duplicated BUSCOs	Fragmented BUSCOs	Missing BUSCOs
D. rosae	1385	277	1108	41	12
[%]	96.3	19.3	77	2.8	0.8
D. rosae* trans.	1388	322	1066	22	28
[%]	96.5	22.3	74	1.5	1.9
M. brunnea	1412	1337	75	20	6
[%]	98.2	93.0	5.2	1.4	0.4
B. cinerea	1423	1340	83	12	3
[%]	98.9	93.2	5.7	0.8	0.2
S. sclerotiorum	1387	1301	86	42	9
[%]	96.5	90.5	5.9	2.9	0.6
M. oryzae	1418	1338	80	10	10
[%]	98.6	93.0	5.5	0.7	0.7
S. pastorianus**	1372	488	884	45	21
[%]	95.4	33.9	61.5	3.1	
A. macrogynus **	1080	316	764	178	180
[%]	75.1	22.0	53.1	12.4	12.5

*predicted transcriptome of *D. rosae*,

**sequences of diploid fungi

https://doi.org/10.1371/journal.pone.0185310.t006

genome were duplicated, whereas other plant pathogen genomes analyzed contain less than 6% duplicated orthologs. Simao *et al*. presented results of their pipeline for 15 fungal genomes and neither of these showed a duplication rate over 11% [58]. This is a strong indication that there is duplication of large parts or the whole genome of *D. rosae*. To examine this hypothesis we tested the genome sequences of two polyploid fungi, the Blastocladiomycete *Allomyces macrogynus* (ATCC 38327) (Broad Institute, Acc. no. ACDU01000000) and the ascomycete *Saccharomyces pastorianus* (CBS 1513) [61] with the BUSCO pipeline. The numbers of duplicated genes are with 53.1% and 61.5% comparable to the results of our genome sequence (Table 6) and support the hypothesis that a genome duplication has occurred relatively recently in *D. rosae*.

Analysis of the duplicated portion of the genome

To rule out technical factors, such as assembly errors, we analyzed the putatively duplicated portion of the black spot genome in more detail. Based on the BUSCO analysis we selected those gene predictions that are BUSCO orthologs and that possess paralogs in the genomic sequence (S8 File). 95.4% of the duplicated BUSCO orthologs occur in pairs, only 3.5% in three copies and less than 1.1% are present at four or five copies in the genome. This is a clear indicator of a single genome duplication. At the same time one would expect a larger diversity of copy numbers if the result was due to an error in the assembly.

Another indicator of sequencing or assembly errors would be a high level of sequence similarity of the duplicated sequences in a haploid genome, because they should represent identical sequences at different locations in the assembly. To determine the degree of polymorphism between the paralogs, we performed a global multiple sequence alignment with the mafft tool [62] on the gene, mRNA and protein level with the predicted sequences and calculated an identity matrix based on these data. Fig 3 illustrates the identity scores of the best pairings (S9 File) as a histogram. The majority of duplications share a high degree of identity (more than 90%), on all three levels in the global alignment. The diversity far exceeds the amount of expected sequencing errors, which could lead to mistakes and duplications in the assembly.

An additional line of evidence comes from the transcriptome data. In more than 97% of the cases both paralogs showed expression for at least one time-point (S8 File). In less than 1% of the cases none of the paralogs shows expression and in 2.3% only one of the duplicates is

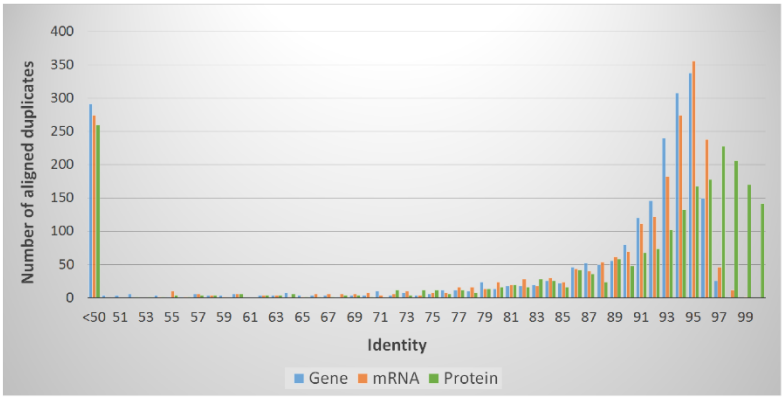


Fig 3. Identity values of *D. rosae* paralogs. Histogram of identity scores generated by global alignment of all duplicated BUSCO orthologs in the *D. rosae* assembly. The gene alignments include intron sequences.

https://doi.org/10.1371/journal.pone.0185310.g003

expressed. If the paralogs were just a product of an erroneous assembly the number of pairs which show expression in only one of the sequences would be much higher because the mapping parameters used were very stringent.

Besides these indictors, which are based on the sequence itself, we developed a PCR-RFLP system to confirm the presence of six paralogous gene pairs (Fig 4A). We designed copy specific PCR primers and tested them on DNA from conidia of the DortE4 isolate which were washed from leaf material, to exclude the possibility that the duplications were due to artificial *in vitro* culture. Differences of restriction sites within the fragments were used as additional evidence that two different gene loci were amplified. For DortE4 all primer pairs produced fragments of the expected sizes and restriction patterns (Fig 4B, Table 7, S10 File). This

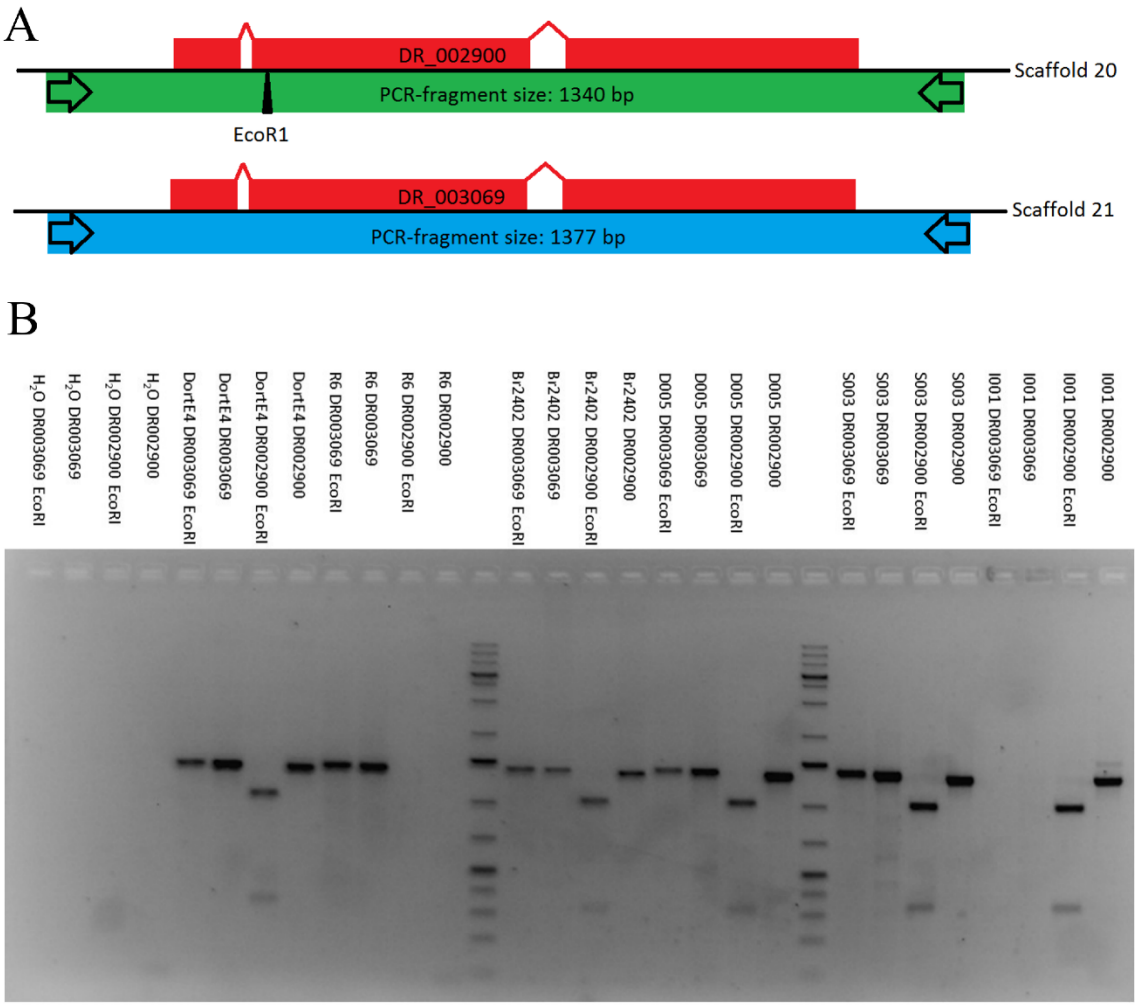


Fig 4. PCR-RFLP analysis of pairs of BUSCO orthologs. A: Schematic display of the method. Specific primers were developed that distinguish between the two very similar paralogs. As second distinctive feature only one of the two PCR fragments contains an EcoRI or DraI restriction site. B: Gel image of a PCR-RFPL analysis of the paralogous sequences DR002900 and DR003069 with six different isolates. Ladder: 1 KB plus (Thermo Fisher, Waltham, USA).

https://doi.org/10.1371/journal.pone.0185310.g004

Table 7. PCR-RFLP results of six pairs of paralogs tested with six different *D. rosae* isolates.

BUSCO-ID	*D. rosae* paralogs	DortE4	R6	I001	S003	D005	Br2402
BUSCOfEOG7D2FR2	DR005870	+	-	+	-	+	+
	DR011096	+	+	-	-	+	+
BUSCOfEOG7K9KD3	DR004273	+	+	-	+	+	+
	DR002774	+	+	+	+	+	+
BUSCOfEOG7MD52Z	DR002315	+	-	+	+	+	+
	DR001421	+	+	-	+	+	+
BUSCOfEOG7R83CB	DR002947	+	+	+	+	+	+
	DR003022	+	+	-	+	+	+
BUSCOfEOG783N6S	DR002900	+	-	+	+	+	+
	DR003069	+	+	-	+	+	+
BUSCOfEOG7DRJFG	DR008457	+	+	-	+	+	+
	DR005710	+	-	-	+	+	-

(+) indicates the presence of the fragment as well as the expected restriction pattern

(-) indicates the absence of a PCR fragment

https://doi.org/10.1371/journal.pone.0185310.t007

confirms the results of the BUSCO analysis and effectively demonstrates experimentally that the genome of the isolate DortE4 contains a large proportion of duplication.

In the next step five additional isolates were tested with the PCR-RFLP system with contrasting results concerning the individual gene duplications (Table 7, Fig 4B). Only D005 contained all duplicates. S003 and Br2402 differed for one of the paralogous pairs. The isolates R6 and I001 show a difference in almost all duplicated pairs but interestingly they do not share the same pattern of differential loss even though they belong to the same race (race seven) [16]. This might indicate that the duplication is already in the state of reduction by gene loss as has been reported for yeast and other fungi [63, 64, 65]. Due to the fact that the tested DNAs originated from conidia isolated from leaves and not from *in vitro* material, the possibility that the inconsistencies in the duplication pattern are the result of stress induced spontaneous duplications can be ruled out. Interestingly, the pattern of gene loss does not appear to be race specific indicating that it might not influence the effector content of fungi but more analyses are needed to clarify that.

The difference between R6 and I001 is also visible in the phylogenetic analysis regarding the paralogous pair DR002900 and DR003069, which were analyzed in more detail with Sanger sequencing of all 10 PCR products (duplicated BUSCO genes from four isolates where I001 and R6 only amplified one of the paralogs). As it can be seen in the maximum likelihood tree (Fig 5, S11 File) all sequences of the two different genes cluster together with the predicted gene models and the corresponding sequences of the genomic scaffold in two distinct clades, proving again the existence of two distinct genes in the *D. rosae* genome where there is only one in the genomes of the related fungi *M. brunnea*, *B. cinerea* and *S. sclerotiorum*. Within these clades I001 and R6 are separated from the other isolates indicating that their sequences contain more polymorphisms than the others.

To get an idea about the extent of the duplication we compared the predicted proteome of *D. rosae* with its relative *M. brunnea* using Blastp. To reduce the influence of gene families we used only those protein sequences of the *M. Brunnea* proteome that do not match other sequences of the proteome than itself (E-value cutoff e-100). These 8402 sequences were used as a reference for a Blastp search (E-value cutoff e-100) with the predicted *D. rosae* proteome and the proteome of *B. cinerea* (S12 File). The number of Blast matches differs between the

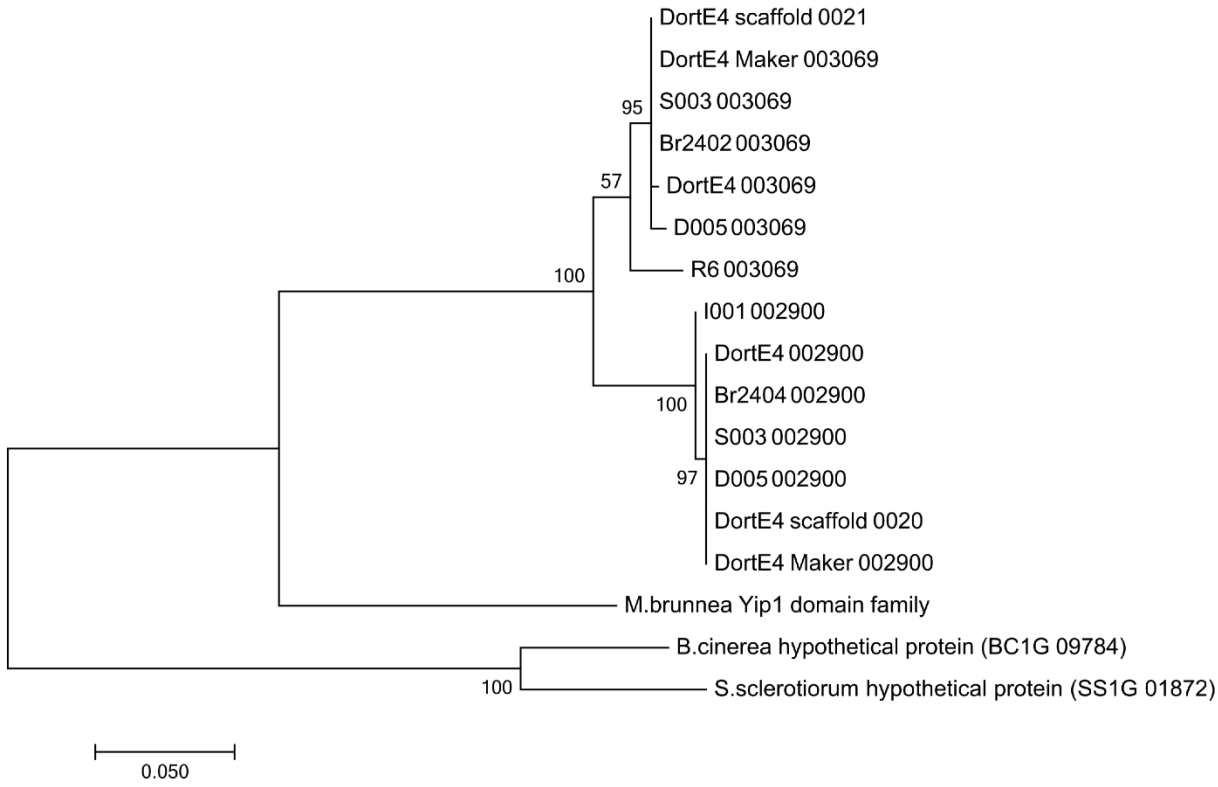


Fig 5. Phylogenetic tree of the paralogous sequences DR002900 and DR003069. The tree contains sequences of six *D. rosae* isolates, the predicted mRNA sequences of the two genes, the genomic region of the scaffolds between the primer binding sites and the corresponding orthologs from *M. brunnea*, *B. cinerea* and *S. sclerotiorum*. The tree was generated with the maximum likelihood method. The values on the branches refer to a bootstrap test (500 replicates).

https://doi.org/10.1371/journal.pone.0185310.g005

species. 5177 of the *M. brunnea* sequences had between one and six matches with sequences from *D. rosea*. Less matches were found between *M. bruunea* and *B. cinerea*. 3724 *M. Brunnea* sequences had up to four matches to the *Botrytis* proteome. This difference is not surprising because the relationship between *D. rosae* and *M. brunnea* is much closer than the relation to *B. cinerea*. All three species belong to the same order but different families. The most important point in this comparison is depicted in Fig 6, where the proportion of blast matches that have different numbers of matches to the same reference sequence is shown. These results indicate that approximately 80% of the *M. brunnea* sequences had two matches to the *D. rosae* proteome and only 16.4% had only one match. The results of the Blast with *B. cinerea* show the complete opposite. Here 93% of the *M. brunnea* sequences had one hit in the *Botrytis* proteome and only a small fraction of 6.3% had two matches. These results correspond with the BUSCO results. Both analyses indicate a single duplication event in which more than three-fourth of the *D. rosae* proteome is duplicated. The amount of single copy genes differ between the two approaches and indicate that 15 to 20% of the genes are still single copy genes (Fig 6).

Overall, the bioinformatic and experimental data indicates that the DortE4 genome contains a large proportion of genome duplication, which is suggestive of a whole genome duplication (WGD) event. The results of the BUSCO analysis and the blast comparison between *D. rose* and *M. brunnea* indicate that the majority of genes are only duplicated once, which argues

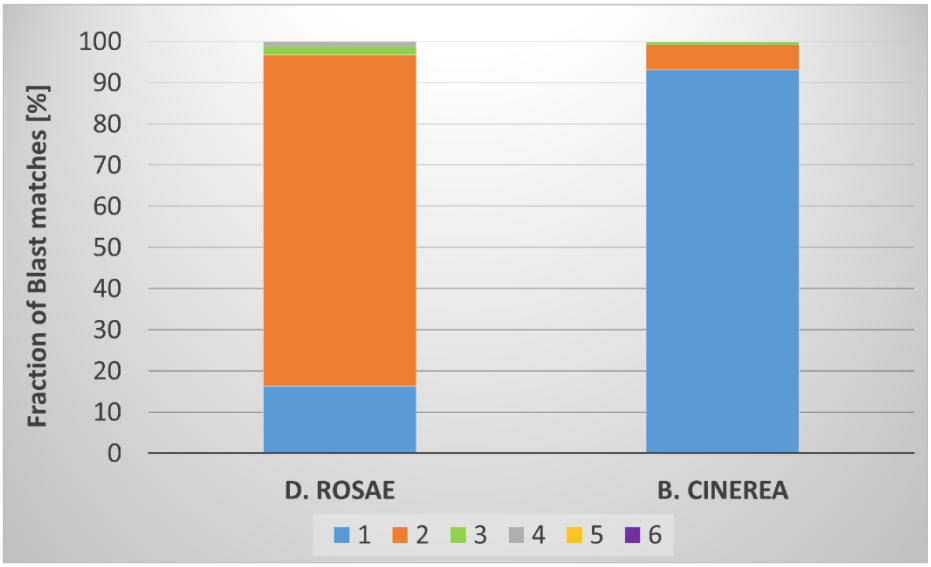


Fig 6. Comparison of the percentage of Blastp matches with different numbers of matches to the same reference sequence between *D. rosae* and *B. cinerea*. The portion of the *M. Brunnea* proteome that had no second match in a Blast (E-value cutoff e-100) of the proteome against itself was the reference for the Blastp search. The cutoff for the Blastp of the related species was set to e-100. The different colors refer to the number of Blast matches to the same *M. Brunnea* reference sequence.

https://doi.org/10.1371/journal.pone.0185310.g006

against multiple segmental duplications and instead for a single duplication event. Based on the high degree of similarity between the different paralogs we can assume that the duplication has occurred relatively recently.

Such duplications are well studied in yeasts and were first described for *Saccharomyces cerevisiae* [66] but have also been shown for many other species [61, 64, 67]. Both allo- and auto-polyploid species are described in the literature.

Reports about WGDs in fungi are not just restricted to yeast, Albertin and Marullo mention in their review of polyploidy in fungi that many other polyploid species occur, e.g. the plant pathogen *Botrytis allii* or the blastocladiomycete *Allomyces macrogynus* mentioned before [66]. In most of the reported cases a WGD event is followed by a loss of duplicated genes [63, 64, 65]. This might be an ongoing process occurring in the *D. rosae* genome and is an explanation for the different patterns of the PCR-RFLP analysis of the tested isolates and the 15–20% of genes that show no duplicated in the BUSCO and Blast analyses. Only a comparison of the complete genome sequence of different isolates can give an answer to this question. Here, the isolates R6 and I001 might be of particular interest due to the clear differences to isolate DortE4. In other fungal species WGDs are often connected to changes in ploidy level. But the unimodal distribution of the k-mer distributions (Fig 1, S1 File) and the relatively low level of similarity between the paralogs (Fig 3) argues against this phenomenon in the *D. rosae* genome because these dissimilarities exceed the amount of variation typically observed for allelic variation.

Gene duplication and gene-loss are key mechanisms in evolution [68, 69, 70]. A WGD event and the following loss of duplicated genes is in most cases a neutral and non-adaptive process, but it can also result in neofunctionalization or subfunctionalization of duplicated

genes, the development of alternative pathways or the remaining of a duplicated gene with redundant function as a genetic buffer [68].

In yeast it is hypothesized that the WGD is a mechanism of diversification, adaption and specialization [66]. In the genome of Microsporidia, which are obligate endoparasitic fungi, it is shown that a WGD event has influenced the local host adaption [71]. For the plant pathogen *Rhizopus oryzae* it was shown that the ancestral WGD has led to an expansion of different gene families that included different virulence factors [65]. It would be interesting to examine if the WGD event influences the *D. rosae* genome in similar way. One hint is the increase in virulence factors. The *D. rosae* genome contains more than four times as many matches in the PHI base than observed for the genome of the closely related pathogen *M. brunnea* [55]. It would be also interesting to analyze how the WGD influences the effector content of the genome. Effector genes are often embedded in highly dynamic regions like TE rich areas or subtelomeric regions, and increase their diversity by rearrangements, duplications, insertions and deletions [6, 72], comparable to the processes occurring during differential gene loss after the genome duplication.

Concluding remarks

We present here the draft genome of *Diplocarpon rosae*, the causative agent of the blackspot disease on roses. With 2457 scaffolds (>500 bp) the assembly is still fragmented. Nevertheless, it contains almost the complete gene space indicated by analysis with the BUSCO pipeline. Noticeable is the estimated genome size ranging from approximately 70 to 90 Mb depending on the used approaches, which is outstandingly large for a fungal genome. This fact cannot be exclusively explained by its content of TEs and other repetitive elements as it is for other fungi. Based on multiple points of evidence, we propose that a whole genome duplication event occurred relatively recently in the genome of *D. rosae*. We could show experimentally that the duplication is neither an artefact of the assembly nor a spontaneous event occurring during *in vitro* culture. There are also indicators that the genome is already in a state of reduction and that different isolates have lost different proportions of the duplicated genome. Whole genome sequence comparison of additional isolates is necessary to clarify this point. Another important question is the influence of the duplication on the pathogenic features of the blackspot fungus. In this regard a more detailed study of the secretome and effector content of the genome is necessary. In conclusion, the *D. rosae* genome sequence is a useful tool to study genome duplication outside of the model system yeast, and in the context of plant pathogenesis. This can give new insights into the evolution of pathogenic features and effector proteins, and will be important in understanding the dynamics of this pathogen as the roles of duplicated loci are analyzed functionally.

Materials and methods

Fungal isolates

Fungal isolates were multiplied on detached leaves of the susceptible rose cultivar "Pariser Charme" as previously published [73]. They comprise the isolates DortE4 (Dortmund, Germany), R6 (Ahrensburg, Germany), Br2402 (Ahrensburg, Germany), I001 (Lucca, Italy), S003 (Sweden), D005 (Groß Lüsewitz, Germany) [16, 24, 73].

Mycelium of the isolate DortE4, which was used for genome sequencing, was also cultivated *in vitro* on 0.8% agar plates containing 3% liquid biomalt extract (Villa Natura, Kirn, Germany).

DNA isolation from mycelium and sequencing

Conidia were washed from agar plates and inoculated into 0.5 l of liquid biomalt medium containing 3% liquid biomalt extract in distilled water (Villa Natura, Kirn, Germany). After six weeks of culture at room temperature with moderate shaking (100 rpm) fungal biomass was harvested by centrifugation. The mycelia were ground in liquid nitrogen in a mortar and DNA was extracted with the Biozym (Hessisch Oldendorf, Germany) MasterPure DNA-extraction kit according to the manufacturer's instructions. Isolated DNA was further purified by one phenol/chloroform extraction and two chloroform extractions, precipitated with isopropanol and subsequently resuspended in TE-buffer and quantified both spectrophotometrically and by gel electrophoresis through 1% agarose gels.

The purified DNA was used for paired-end library preparations with the Epicentre Nextera DNA Library Preparation Kit (Epicentre, Madison, USA) according to the manufacturer's guidelines. After preparation the library was size selected on a 1% agarose gel and purified using the Qiagen Gel Extraction kit (Hilden, Germany) to produce an average insert length of approximately 500bp. A mate-pair library was prepared with the same DNA using the Nextera Mate-Pair Library Preparation Kit (Illumina, San Diego, USA)) with size-selection (~3 kb) performed on a 0,75% agarose gel and purified using the Qiagen Gel Extraction Kit. The paired-end library was sequenced on an Illumina HiScan SQ with 2x100bp (v3) chemistry. The mate-pair library was sequenced on an Illumina HiSeq 2500 with 2x125bp read length (v4) chemistry.

454 sequence data was generated by Macrogen (Seoul, South Korea), who performed the library preparation and pyrosequencing on a Roche (Basel, Switzerland) 454 GS-FLX system.

Trimming and assembly

Adapter and quality trimming was performed with the CLC Genomics Workbench (Qiagen, Hilden, Germany). Removal of adaptor content and trimming of poor quality data (<Q20) of the 454 reads was performed with the CLC Genomics Workbench (Qiagen, Hilden, Germany) resulting in an average read length of 440 bp.

The same software was also used for trimming of the Illumina paired-end sequences of the short insert library to a read quality below a PHRED score of Q20 and for trimming of the 3' ends to a read length of 2x80bp.

Mate-pair data were trimmed using Trimmomatic 0.32 [74] to remove Nextera transposase and TruSeq adaptor sequences as well as poor quality data (<Q20).

The *de novo* assembly algorithm of CLC Genomics Workbench 6.5 (Qiagen, Hilden, Germany) was used to generate a contig-level *de novo* assembly using a combination of short insert library paired-end Illumina and 454 data. The assembly was performed with a "word size" (k-mer size) of 50 and automatic "bubble size". Contigs were updated after re-mapping reads to the initial assembly and a minimum contig size of 200bp was selected. Scaffolding with paired-end data was performed as part of the CLC *de novo* assembly. Further scaffolding with mate-pair data was performed using Opera-LG (v2) [75]. For Opera scaffolding a ploidy level of 1 was specified and mapping of mate-pair reads to contigs was performed with the BWA software package (version 0.7) [76].

Gene prediction and annotation

For the gene prediction and structural annotation the MAKER 2.3.1.8 pipeline [35] was used, which combines repeat masking, different prediction tools with evidence based quality control and gene model editing. The repeat masking was done by using RepeatMasker 4.0 [77] with RMBlast search algorithm and the Repbase [78] database. Three different *de novo* prediction tools were combined in the pipeline: the self-training tool GeneMark-ES [37], Augustus [36]

with the prediction models of *B. cinerea* and SNAP 3 [79] which was trained by three rounds of hint-based MAKER prediction. As evidence for MAKER annotation the predicted proteome of the closely related fungus *M. brunnea* [55] and an assembly of the RNA-Seq and MACE data were used. As an additional tool tRNAscan [80] was integrated.

The functional annotation was done with Blast2GO 3.3 [40] and Blastx against the NCBI NR protein database (E-value cutoff e-10). The integrated Gene Ontology (GO) term annotation [81] and InterProScan [41] were both performed using default parameters.

Additional bioinformatic analysis

To determine the genome size, the tool Jellyfish 1.1.10 [82] was used to produce a k-mer distribution (k-mer size 17, 20, 22, 25, 30, 35) with reads from a small insert library. The data processing was done in R [83]. The genome size was calculated according to procedure used by Li and colleagues for the giant panda sequence [29] and by Liu *et al.* used for the *Brassica oleracea* genome [30] and the GenomScope Software [31] (S1 File).

The repeat structure of the genome was analyzed with the RepeatMasker 4.0.6 software [77] in combination with the RepeatModeler package 1.0.8 [84] and the Repbase [78] database. The cross-match search engine from the Phrap package [85] was applied for this screening. For a more detailed screening for SSR-motifs, the SSR Locator software [56] was applied with a minimum of ten repeats per motif type.

Analyses with the BUSCO pipeline version 1.2 [58] were performed, in genome as well as in transcriptome mode with the fungi dataset.

Different variants of the BLAST+ package [86] were used for all vs all Blast alignments of the paralogous sequences with an E-value cutoff of 1e-10.

For calculation the identity scores of the duplicated BUSCO orthologs a multiple sequence alignment was performed with the mafft 7.3 tool using the FFT-NS-1 alignment algorithm [62]. The resulting alignment was used for generating an identity matrix with the software trimAl 1.2 [87]

DNA isolation from conidia and PCR-RFLP of paralogous sequences

DNA for this analysis was isolated from conidia that were washed from infected "Pariser Charme" leaves according to the manufacturer's instructions with the MasterPure DNA-extraction kit (Biozym, Hessisch Oldendorf, Germany). The only divergence from the protocol was the extension of all incubation steps.

The PCR-RFLP analysis was performed for six BUSCO orthologous pairs and six *D. rosae* isolates. To distinguish between the two paralogous sequences, specific primer pairs were designed with the Primer3 web tool [88]. PCR reactions were performed in a 20 μl volume containing 10 ng DNA, 0.1 mM dNTP, 0.25 μM forward and reverse primer, 1 U *Taq* polymerase (DCS, DNA cloning service, Hamburg, Germany) and 1x reaction buffer [89]. The following cycling conditions were used: initial denaturation for 10 min at 94˚C, followed by 30–35 cycles of 60s at 94˚C, 60s at 64˚C and 120s at 72˚C and a final elongation of 10 min at 72˚C.

The PCR reaction was then used for restriction digestion with EcoRI or DraI (Thermo Fisher Scientific, Waltham, USA). Both PCR and restriction digestion products were separated on 1.5% agarose gels.

The amplified fragments of the genes DR003069 and DR002900 of all isolates were sequenced with the Sanger technique. The PCR was performed with the Primestar proof reading *Taq* polymerase (Takara Bio, Clontech, Mountain View, USA) and the above-mentioned amplification protocol. Only the cycling conditions were changed as followed: 98˚C for 5 min, 32 cycles of 10 s at 98˚C, 15 s at 64˚C and 120s at 72˚C following the final elongation of 10 min

at 72˚C. PCR-products were purified with the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel, Düren, Germany) following the manufacturer's instructions. Fragments were ligated into the pJET cloning vector (Thermo Fisher Waltham, USA) using the blunt-end ligation protocol. Sanger sequencing was performed by GATC Biotech (Konstanz, Germany) using the standard pJET sequencing primer.

The resulting sequences from the isolates and the corresponding gene sequences of related fungi were used for generation of a phylogenetic tree using the software MEGA 7.0.18 [90]. The tree was constructed with the Maximum Likelihood method based on the Tamura-Nei model [91]. The percentage of replicate trees in which the associated sequences clustered together in the bootstrap test (500 replicates) are shown next to the branches [92].

Generation of the expression data

MACE (Massive Analysis of cDNA Ends) data [43, 93] for three time points (0, 24, 72 hpi) and three biological replicates for the compatible interaction of the isolate DortE4 with the susceptible rose variety "Pariser Charme" (PC) were generated so that each of the three biological repeats is derived from an independent inoculation experiment. To expand the analysis, the RNA of the three samples of the 72 hpi time point were also used for a conventional RNA-Seq approach. PC leaves were spray inoculated in a detached leaf assay with a spore concentration of $5x10^5$ spores per ml.

Immediately after sampling, 30 mg leaf material was frozen in liquid nitrogen and RNA was extracted with the Qiagen (Hilden, Germany) RNeasy plant mini kit according to manufacturer's manual. Library preparation and sequencing was either done by GenXPro (Frankfurt am Main, Germany) for the MACE data or GATC biotech (Konstanz, Germany) for the RNA-Seq data.

Data analyses were performed with the CLC Genomics Workbench 9.0.1 (Qiagen, Hilden, Germany). Quality trimming was done with default parameters. Read mapping for expression profiling was performed with default parameters except the similarity and length fraction which were set to 0.9 respectively 0.95 to increase the sensitivity. A gene was considered to be expressed at a time point if reads were mapped in a least two repeated experiments. Quantitative gene expression levels were not taken into account due to the increasing biomass of the fungus, instead the genes were only classified as expressed or not expressed.

Supporting information

S1 File. Genome size estimation by the k-mer distribution.
(DOCX)

S2 File. Annotation statistics.
(DOCX)

S3 File. Blast2Go annotation table.
(XLSX)

S4 File. Expression data generated with MACE and RNA-Seq.
(XLSX)

S5 File. Result of the Blast to the Phi database.
(XLSX)

S6 File. CAZy-annotation.
(XLSX)

S7 File. RepeatMasker results.
(XLSX)

S8 File. Results of the BUSCO pipeline.
(XLSX)

S9 File. Identity values of the duplicated BUSCO orthologs.
(XLSX)

S10 File. Details about the PCR-RFLP.
(DOCX)

S11 File. Alignment of the paralogous sequences DR002900 and DR003069.
(FAS)

S12 File. BlastP to the *M. brunnea* proteom.
(XLSX)

Acknowledgments

We give special thanks to Dr. Helgard Kaufmann who was involved in the generation of the MACE data. Even if these particular datasets were not included in this study the transcriptomic analysis would have been impossible without her contribution.

Author Contributions

Conceptualization: Enzo Neu, Jonathan Featherston, Jasper Rees, Thomas Debener.

Data curation: Enzo Neu, Jonathan Featherston, Thomas Debener.

Funding acquisition: Jasper Rees, Thomas Debener.

Investigation: Enzo Neu, Jonathan Featherston, Jasper Rees, Thomas Debener.

Methodology: Enzo Neu, Jonathan Featherston, Jasper Rees, Thomas Debener.

Project administration: Jasper Rees, Thomas Debener.

Resources: Enzo Neu, Jonathan Featherston, Jasper Rees, Thomas Debener.

Supervision: Jasper Rees, Thomas Debener.

Validation: Enzo Neu, Jonathan Featherston, Jasper Rees, Thomas Debener.

Visualization: Enzo Neu, Jonathan Featherston, Jasper Rees, Thomas Debener.

Writing – original draft: Enzo Neu, Jonathan Featherston, Jasper Rees, Thomas Debener.

Writing – review & editing: Enzo Neu, Jonathan Featherston, Jasper Rees, Thomas Debener.

References

1. Boyd LA, Ridout C, O'Sullivan DM, Leach JE, Leung H. Plant-pathogen interactions: disease resistance in modern agriculture. Trends in Genetics. 2013; 29: 233–240. https://doi.org/10.1016/j.tig.2012.10.011 PMID: 23153595
2. Collinge DB, Jorgensen HJL, Lund OS, Lyngkjaer MF. Engineering pathogen resistance in crop plants: current trends and future prospects. Annual Review of Phytopathology. 2010; 48: 269–291. https://doi.org/10.1146/annurev-phyto-073009-114430 PMID: 20687833
3. Cui H, Tsuda K, Parker JE. Effector-triggered immunity: from pathogen perception to robust defense. Annual Review of Plant Biology. 2015; 66: 487–511. https://doi.org/10.1146/annurev-arplant-050213-040012 PMID: 25494461

4. Dodds PN, Rathjen JP. Plant immunity: towards an integrated view of plant-pathogen interactions. Nature Reviews Genetics. 2010; 11: 539–548. https://doi.org/10.1038/nrg2812 PMID: 20585331
5. Pedro H, Maheswari U, Urban M, Irvine AG, Cuzick A, McDowall MD, et al. PhytoPath: an integrative resource for plant pathogen genomics. Nucleic Acids Research. 2016; 44: D688–93. https://doi.org/10.1093/nar/gkv1052 PMID: 26476449
6. Selin C, de Kievit TR, Belmonte MF, Fernando WGD. Elucidating the Role of Effectors in Plant-Fungal Interactions: Progress and Challenges. Frontiers in Microbiology. 2016; 7(600).
7. Sharma KK. Fungal genome sequencing: basic biology to biotechnology. Critical Reviews in Biotechnology. 2016; 36: 743–759. https://doi.org/10.3109/07388551.2015.1015959 PMID: 25721271
8. Shirke MD, Mahesh HB, Gowda M. Genome-Wide Comparison of Magnaporthe Species Reveals a Host-Specific Pattern of Secretory Proteins and Transposable Elements. PLOS ONE. 2016; 11(9).
9. Chand R, Pal C, Singh V, Kumar M, Singh VK, Chowdappa P. Draft genome sequence of Cercospora canescens: a leaf spot causing pathogen. Current Science. 2015; 109(11).
10. Aragona M, Minio A, Ferrarini A, Valente MT, Bagnaresi P, Orru L, et al. De novo genome assembly of the soil-borne fungus and tomato pathogen Pyrenochaeta lycopersici. BMC Genomics. 2014; 15(1).
11. Kubicek CP, Starr TL, Glass NL. Plant Cell Wall-Degrading Enzymes and Their Secretion in Plant-Pathogenic Fungi. Annual Review of Phytopathology 2014, 52:427–451. https://doi.org/10.1146/annurev-phyto-102313-045831 PMID: 25001456
12. Lo Presti L, Lanver D, Schweizer G, Tanaka S, Liang L, Tollot M, et al. Fungal Effectors and Plant Susceptibility. Annual Review of Plant Biology. 2015; 66: 513–545. https://doi.org/10.1146/annurev-arplant-043014-114623 PMID: 25923844
13. Horst RK, Cloyd RA. Compendium of rose diseases and pests. 2nd ed. St. Paul, Minn., USA: APS Press; 2007.
14. Gachomo EW, Dehne H, Steiner U. Microscopic evidence for the hemibiotrophic nature of Diplocarpon rosae, cause of black spot disease of rose. Physiological and Molecular Plant Pathology. 2006; 69: 86–92.
15. Debener T, Byrne DH. Disease resistance breeding in rose: current status and potential of biotechnological tools. Plant Science. 2014; 228: 107–117. https://doi.org/10.1016/j.plantsci.2014.04.005 PMID: 25438791
16. Whitaker VM, Debener T, Roberts AV, Hokanson SC. A standard set of host differentials and unified nomenclature for an international collection of Diplocarpon rosae races. Plant Pathology. 2010; 59: 745–752.
17. Blechert O, Debener T. Morphological characterization of the interaction between Diplocarpon rosae and various rose species. Plant Pathology. 2005; 54: 82–90.
18. Gachomo EW, Seufferheld MJ, Kotchoni SO. Melanization of appressoria is critical for the pathogenicity of Diplocarpon rosae. Molecular Biology Reports. 2010; 37: 3583–3591. https://doi.org/10.1007/s11033-010-0007-4 PMID: 20204524
19. von Malek B, Debener T. Genetic analysis of resistance to blackspot (Diplocarpon rosae) in tetraploid roses. Theoretical and Applied Genetics. 1998; 96: 228–231.
20. Hattendorf A, Linde M, Mattiesch L, Debener T, Kaufmann H. Genetic Analysis of Rose Resistance Genes and their Localisation in the Rose Genome. Acta Horticulturae. 2004: 123–130.
21. Whitaker VM, Bradeen JM, Debener T, Biber A, Hokanson SC. Rdr3, a novel locus conferring black spot disease resistance in tetraploid rose: genetic analysis, LRR profiling, and SCAR marker development Theoretical and Applied Genetics. 2010; 120: 573–585. https://doi.org/10.1007/s00122-009-1177-0 PMID: 19847388
22. Kaufmann H, Mattiesch L, Lorz H, Debener T. Construction of a BAC library of Rosa rugosa Thunb. and assembly of a contig spanning Rdr1, a gene that confers resistance to blackspot. Molecular Genetics and Genomics. 2003; 268: 666–674. https://doi.org/10.1007/s00438-002-0784-0 PMID: 12589441
23. Terefe-Ayana D, Yasmin A, Le TL, Kaufmann H, Biber A, Kuhr A, et al. Mining disease-resistance genes in roses: functional and molecular characterization of the rdr1 locus. Frontiers in Plant Science. 2011; 2:35. https://doi.org/10.3389/fpls.2011.00035 PMID: 22639591
24. Menz I, Straube J, Linde M, Debener T. The TNL gene Rdr1 confers broad-spectrum resistance to Diplocarpon rosae. Molecular Plant Pathology. 2017
25. Münnekhoff AK, Linde M, Debener T: The gene diversity pattern of *Diplocarpon rosae* populations is shaped by the age, diversity and fungicide treatment of their host populations. Plant Pathology. 2017.
26. Bock CH, Chen C, Yu F, Stevenson KL, Wood BW. Draft genome sequence of Fusicladium effusum, cause of pecan scab. Standards in Genomic Sciences. 2016; 11(36).

27. D'Hondt L, Hofte M, van Bockstaele E, Leus L. Applications of flow cytometry in plant pathology for genome size determination, detection and physiological status. Molecular Plant Pathology. 2011; 12: 815–828. https://doi.org/10.1111/j.1364-3703.2011.00711.x PMID: 21726378

28. Veselska T, Svoboda J, Ruzickova Z, Kolarik M. Application of Flow Cytometry for Genome Size Determination in Geosmithia Fungi: A Comparison of Methods. Cytometry Part A. 2014; 85A: 854–861.

29. Li R, Fan W, Tian G, Zhu H, He L, Cai J, et al. The sequence and de novo assembly of the giant panda genome. Nature. 2010; 463: 311–317. https://doi.org/10.1038/nature08696 PMID: 20010809

30. Liu S, Liu Y, Yang X, Tong C, Edwards D, Parkin I A P, et al. The Brassica oleracea genome reveals the asymmetrical evolution of polyploid genomes. Nature Communications. 2014; 5(3930).

31. Vurture G, Sedlazeck F, Nattestad M, Underwood C, Fang H, Gurtowski J, Schatz M. GenomeScope: fast reference-free genome profiling from short reads. Bioinformatics. 2017; 33(14): 2202–2204. https://doi.org/10.1093/bioinformatics/btx153 PMID: 28369201

32. Gregory TR, Nicol JA, Tamm H, Kullman B, Kullman K, Leitch IJ, et al. Eukaryotic genome size databases. Nucleic Acids Research. 2007; 35: D332–338. https://doi.org/10.1093/nar/gkl828 PMID: 17090588

33. Fraissinet-Tachet L, Reymond-Cotton P, Fèvre M. Molecular karyotype of the phytopathogenic fungus-Sclerotinia sclerotiorum. Current genetics. 1996; 29(5): 496–501 PMID: 8625431

34. Weber E. Untersuchungen zu Fortpflanzung und Ploidie verschiedener Ascomyceten. Bibliotheca Mycologica. 1992; 140: 1–186.

35. Cantarel BL, Korf I, Robb SM C, Parra G, Ross E, Moore B, et al. MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Research. 2008; 18: 188–196. https://doi.org/10.1101/gr.6743907 PMID: 18025269

36. Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. AUGUSTUS: ab initio prediction of alternative transcripts. Nuclei Acids Recearch. 2006; 34: W435–W439.35.

37. Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. Genome Research. 2008; 18: 1979–1990. https://doi.org/10.1101/gr.081612.108 PMID: 18757608

38. Galagan JE, Henn, Ma LJ, Cuomo CA, Birren B. Genomics of the fungal kingdom: Insights into eukaryotic biology. Genome Research. 2005; 15: 1620–1631. https://doi.org/10.1101/gr.3767105 PMID: 16339359

39. Mohanta TK, Bae H. The diversity of fungal genome. Biological Procedures Online. 2015; 17(8).

40. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics. 2005; 21: 3674–3676. https://doi.org/10.1093/bioinformatics/bti610 PMID: 16081474

41. Jones P, Binns D, Chang H, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014; 30: 1236–1240. https://doi.org/10.1093/bioinformatics/btu031 PMID: 24451626

42. Bairoch A. The ENZYME database in 2000. Nucleic Acids Research. 2000; 28: 304–305. PMID: 10592255

43. Zawada AM, Rogacev KS, Muller S, Rotter B, Winter P, Fliser D,et al.: Massive analysis of cDNA Ends (MACE) and miRNA expression profiling identifies proatherogenic pathways in chronic kidney disease. Epigenetics. 2014; 9: 161–172. https://doi.org/10.4161/epi.26931 PMID: 24184689

44. Winnenburg R, Baldwin TK, Urban M, Rawlings C, Koehler J, Hammond-Kosack KE. PHI-base: a new database for pathogen host interactions. Nucleic Acids Research. 2006; 34: D459–D464. https://doi.org/10.1093/nar/gkj047 PMID: 16381911

45. Winnenburg R, Urban M, Beacham A, Baldwin TK, Holland S, Lindeberg M, et al. PHI-base update: additions to the pathogen-host interaction database. Nucleic Acids Research. 2008; 36: D572–D576. https://doi.org/10.1093/nar/gkm858 PMID: 17942425

46. Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Research. 2014; 42: D490–D495. https://doi.org/10.1093/nar/gkt1178 PMID: 24270786

47. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Research. 2012; 40: W445–W451. https://doi.org/10.1093/nar/gks479 PMID: 22645317

48. Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. Biotechnology for Biofuels. 2013; 6(41).

49. de Jonge R, van Esse HP, Kombrink A, Shinya T, Desaki Y, Bours R, et al. Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. Science. 2010; 329: 953–955. https://doi.org/10.1126/science.1190859 PMID: 20724636

50. Jiang C, He B, Huang R, Huang M, Xu L. Identification and functional analysis of LysM effectors from Marssonina brunnea. Australasian Plan Pathology. 2014; 43: 615–622.

51. Grandaubert J, Balesdent M, Rouxel T. Evolutionary and Adaptive Role of Transposable Elements in Fungal Genomes. Advances in Botanical Research. 2014; 70: 79–107.

52. Stukenbrock EH, Croll. The evolving fungal genome. Fungal Biology Reviews. 2014; 28: 1–12.

53. Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, Stüber K, et al. Genome Expansion and Gene Loss in Powdery Mildew Fungi Reveal Tradeoffs in Extreme Parasitism. Science. 2010; 330: 1543–1546 https://doi.org/10.1126/science.1194573 PMID: 21148392

54. Martin F, Kohler A, Murat C, Balestrini R, Coutinho PM, Jaillon O, et al. Perigord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. Nature. 2010; 464: 1033–1038. https://doi.org/10.1038/nature08867 PMID: 20348908

55. Zhu S, Cao Y, Jiang C, Tan B, Wang Z, Feng S, et al. Sequencing the genome of Marssonina brunnea reveals fungus-poplar co-evolution. BMC Genomics. 2012; 13(382).

56. da Maia LC, Palmieri DA, de Souza VQ, Kopp MM, de Carvalho FIF, de Costa OA. SSR Locator: Tool for Simple Sequence Repeat Discovery Integrated with Primer Design and PCR Simulation. International Journal of Plant Genomics. 2008.

57. Karaoglu H, Lee CM, Meyer W. Survey of simple sequence repeats in completed fungal genomes. Molecular Biology and Evolution. 2005; 22: 639–649. https://doi.org/10.1093/molbev/msi057 PMID: 15563717

58. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015; 31: 3210–3212. https://doi.org/10.1093/bioinformatics/btv351 PMID: 26059717

59. Amselem J, Cuomo CA, van Kan JA L, Viaud M, Benito EP, Couloux A, et al. Genomic Analysis of the Necrotrophic Fungal Pathogens Sclerotinia sclerotiorum and Botrytis cinerea. PLOS Genetics. 2011; 7 (8).

60. Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, et al. The genome sequence of the rice blast fungus Magnaporthe grisea. Nature. 2005; 434: 980–986. https://doi.org/10.1038/nature03449 PMID: 15846337

61. Okuno M, Kajitani R, Ryusui R, Morimoto H, Kodama Y, Itoh T. Next-generation sequencing analysis of lager brewing yeast strains reveals the evolutionary history of interspecies hybridization. DNA Research. 2016; 23: 67–80. https://doi.org/10.1093/dnares/dsv037 PMID: 26732986

62. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution. 2013; 16: 772–780.

63. Scannell, Byrne KP, Gordon JL, Wong S, Wolfe KH. Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. Nature. 2006; 440: 341–345. https://doi.org/10.1038/nature04562 PMID: 16541074

64. Scannell DR, Frank AC, Conant GC, Byrne KP, Woolfit M, Wolfe KH. Independent sorting-out of thousands of duplicated gene pairs in two yeast species descended from a whole-genome duplication. Proceedings of the National Academy of Sciences of the United States of America PNAS. 2007; 104: 8397–8402. https://doi.org/10.1073/pnas.0608218104 PMID: 17494770

65. Ma L, Ibrahim AS, Skory C, Grabherr MG, Burger G, Butler M, et al. Genomic Analysis of the Basal Lineage Fungus Rhizopus oryzae Reveals a Whole-Genome Duplication. PLOS Genetics. 2009; 5(7).

66. Albertin W, Marullo P. Polyploidy in fungi: evolution after whole-genome duplication. Proceedings of the Royal Society -Biological Sciences. 2012; 279: 2497–2509. https://doi.org/10.1098/rspb.2012.0434 PMID: 22492065

67. Naumov GI, Naumova ES, Masneuf I, Aigle M, Kondratieva VI, Dubourdieu D. Natural Polyploidization of Some Cultured Yeast Saccharomyces Sensu Stricto. Auto- and Allotetraploidy. Systematic and Applied Microbiology. 2000; 23: 442–449. https://doi.org/10.1016/S0723-2020(00)80076-4 PMID: 11108025

68. Albalat R, Cañestro C. Evolution by gene loss. Nature Reviews Genetics. 2016; 17(7): 379–391. https://doi.org/10.1038/nrg.2016.39 PMID: 27087500

69. del Pozo JC, Ramirez-Parra E. Whole genome duplications in plants: an overview from Arabidopsis. Journal of Experimental Botany. 2015; 66(22): 6991–7003. https://doi.org/10.1093/jxb/erv432 PMID: 26417017

70. Panchy N, Lehti-Shiu M, Shiu SH. Evolution of Gene Duplication in Plants. Plant Physiology. 2016; 171: 2294–2316. https://doi.org/10.1104/pp.16.00523 PMID: 27288366

71. Williams TA, Nakjang S, Campbell SA, Freeman MA, Eydal M, Moore K, et al. A Recent Whole-Genome Duplication Divides Populations of a Globally Distributed Microsporidian. Molecular Biology and Evolution. 2016; 33(8): 2002–2015. https://doi.org/10.1093/molbev/msw083 PMID: 27189558

72. de Wit P, Mehrabi R, van den Burg H, Stergiopoulos I. Fungal effector proteins: past, present and future. Molecular Plant Pathology. 2009; 10: 735–747. https://doi.org/10.1111/j.1364-3703.2009.00591.x PMID: 19849781

73. Debener T, Drewes Alvarez R, Rockstroh K. Identification of five physiological races of blackspot, Diplocarpon rosas, Wolf on roses. Plant Breeding. 1998; 117: 267–270.

74. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30: 2114–2120. https://doi.org/10.1093/bioinformatics/btu170 PMID: 24695404

75. Gao S, Bertrand D, Chia BKH, Nagarajan N. OPERA-LG: efficient and exact scaffolding of large, repeat-rich eukaryotic genomes with performance guarantees. Genome Biology. 2016; 17(102).

76. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25: 1754–1760. https://doi.org/10.1093/bioinformatics/btp324 PMID: 19451168

77. Smit AFA, Hubley R, Green P. RepeatMasker Open-4.0; http://www.repeatmasker.org/. Accessed 16 Dec. 2016

78. Bao W, Kojima KK, Kohany O. Repbase Update, a database of repetitive elements in eukaryotic genomes. Mobile DNA. 2015; 6(11).

79. Korf I. Gene finding in novel genomes. BMC Bioinformatics. 2004; 5(59).

80. Lowe TM, Eddy SR. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. Nuclei Acids Research. 1997; 25: 955–964.

81. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology Consortium: Gene Ontology: tool for the unification of biology. Nature Genetics. 2000; 25: 25–29. https://doi.org/10.1038/75556 PMID: 10802651

82. Marcais G, Kingsford C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics. 2011; 27: 764–770. https://doi.org/10.1093/bioinformatics/btr011 PMID: 21217122

83. Team RC. R: A Language and Environment for Statistical Computing [https://www.R-project.org/].

84. Smit AFA, Hubley R. RepeatModeler Open-1.0; http://www.repeatmasker.org/. Accessed 16 Dec. 2016

85. Green P, Brent E. phred 0.020425.c, http://www.phrap.org/phredphrapconsed.html Accessed 16 Dec. 201674.

86. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009; 10(421).

87. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyse. 2009 Bioinformatics; 25: 1972–1973 https://doi.org/10.1093/bioinformatics/btp348 PMID: 19505945

88. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3-new capabilities and interfaces. Nucleic Acids Research 2012, 40(15):e115. https://doi.org/10.1093/nar/gks596 PMID: 22730293

89. Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research. 1990; 18: 6531–6535. PMID: 1979162

90. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular Biology and Evolution. 2016; 33: 1870–1874. https://doi.org/10.1093/molbev/msw054 PMID: 27004904

91. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution. 1993; 10: 512–526. PMID: 8336541

92. Felsenstein J. Confidence Limits on Phylogenies. An Approach Using the Bootstrap. Evolution. 1985; 39: 783–791. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x PMID: 28561359

93. Zajac BK, Amendt J, Horres R, Verhoff MA, Zehner R. De novo transcriptome analysis and highly sensitive digital gene expression profiling of Calliphora vicina (Diptera: Calliphoridae) pupae using MACE (Massive Analysis of cDNA Ends). Forensic Science International Genetics. 2015; 15: 137–146. https://doi.org/10.1016/j.fsigen.2014.11.013 PMID: 25468442

4. Prediction of the *Diplocarpon rosae* secretome reveals candidate genes for effectors and virulence factors

Enzo Neu[1], Thomas Debener[1]*

[1.] Department of Molecular Plant Breeding, Institute for Plant Genetics, Leibniz Universität Hannover, Hannover 30419, Germany

Type of authorship:	First author
Type of article:	Research article
Contribution to the article:	Planned and performed all analysis Wrote the paper Prepared all tables and figures
Contribution of the other authors:	Thomas Debener contributed in planning of the analysis, as well as in the revision of the manuscript and writing of parts of the introduction
Journal:	Fungal Biology
Date of submission:	25.04.2018
Impact factor:	2.184

Prediction of the *Diplocarpon rosae* secretome reveals candidate genes for effectors and virulence factors

Enzo Neu[1], Thomas Debener[1]*

[1.] Department of Molecular Plant Breeding, Institute for Plant Genetics, Leibniz Universität Hannover, Hannover 30419, Germany

* debener@genetik.uni-hannover.de

Research highlights

- Prediction of 827 secreted proteins by a combination of different bioinformatic tools
- A significant proportion of the secretome comprises of plant CWDEs
- 52 of the 251 effector candidates match several bioinformatic criteria
- These effector candidates also contain a Y/F/WxC motif of biotrophic fungi
- Several candidates are highly expressed during penetration/haustorium formation

Keywords: fungi, secretome, effector prediction, black spot, roses, plant-pathogen interaction, plant cell wall-degrading enzymes

Total word count: 6809

Word count Abstract: 226

Abstract

Rose black spot is one of the most severe diseases of field-grown roses. Though R-genes have been mapped and one TNL gene has been isolated, little information is known about the molecular details of the interaction between pathogen and host. Based on the recently published genome sequence of the black spot fungus, we analysed gene models with various bioinformatic tools utilising the expression data of infected host tissues, which led to the prediction of 827 secreted proteins. A significant proportion of the predicted secretome comprises enzymes for the degradation of cell wall components, several of which were highly expressed during the first infection stages. As the secretome comprises major factors determining the ability of the fungus to colonise its host, we focused our further analyses on predicted effector candidates. In total, 52 sequences of 251 effector candidates matched several bioinformatic criteria of effectors, contained a Y/F/WxC motif, and did not match annotated proteins from other fungi. Additional sequences were identified based on their high expression levels during the penetration/haustorium formation phase and/or by matching known effectors from other fungi. Several host genotypes that are resistant to the sequenced isolate but differ in the R-genes responsible for this resistance are available. The combination of these genotypes with functional studies of the identified candidate effectors will allow the mechanisms of the rose black spot interaction to be dissected.

1. Introduction

Black spot disease is one of the most severe and damaging diseases of field-grown roses. It is caused by the hemibiotrophic ascomycete *Diplocarpon rosae* (whose anamorph is *Marssonina rosae*). Infections with this pathogen lead to chlorotic and necrotic spots on the leaves, drastically reducing the ornamental value of these plants. In the case of stronger infections, the disease can even lead to the defoliation or death of highly susceptible genotypes. The fungus spreads mainly though asexual conidia by splash water or direct contact (Horst and Cloyd 2007). After the germination of bicellular conidia, the fungus penetrates the cuticle via an appressorium within the first 12 hours post-inoculation (hpi) and develops intercellular haustoria to extract nutrients from the plant. This process takes two to three days and

marks the early biotrophic stage of the pathogen, which is followed by a mixed biotrophic/necrotrophic stage that leads to some tissue damage and results (after five to seven days) in the development of acervuli, where new conidia are formed (Aronescu 1934; Frick 1943; Gachomo and Kotchoni 2007).

Due to the high economic importance of roses and the severe damage the fungus causes to them, the interaction of this particular fungus with cultivated roses is one of the best studied (Debener and Byrne 2014). Many single spore isolates are available, and various authors have used different sets of host plants to differentiate up to 11 pathogenic races (Debener *et al.* 1998; Whitaker *et al.* 2007; Whitaker *et al.* 2010b). The genetic diversity of *D. rosae* has also been studied with different types of molecular markers (Werlemark *et al.* 2006; Whitaker *et al.* 2007; Münnekhoff *et al.* 2017), and its interaction with its host has been studied by histological and biochemical methods (Blechert and Debener 2005; Gachomo *et al.* 2006; Gachomo *et al.* 2010). The main research goal on the host side is to gain resistance against *D. rosae*, and several R-loci (resistance loci) are already characterised (Malek and Debener 1998; Hattendorf *et al.* 2004; Whitaker *et al.* 2010a). The best-studied locus is Rdr1, which mediates broad-spectrum resistance against different races and is encoded by a TNL-type resistance gene that recognises an effector of *D. rosae*, which leads to a defence response (Kaufmann *et al.* 2003; Terefe-Ayana *et al.* 2011; Menz *et al.* 2017).

The fungal secretome is defined as the entirety of proteins that are secreted outside the plasma membrane of a cell (Girard *et al.* 2013; Meinken *et al.* 2016). This definition differs slightly from the original definition made for *Bacillus subtilis*, where an analysis of the secretory pathway machinery was also included in the secretome (Tjalsma *et al.* 2000). Most secreted proteins carry a so-called signal peptide (SP), a signal sequence at the N-terminus containing three parts: a positively charged n-region, a neutral or polar c-region and a hydrophobic region of approximately 5-16 amino acids with a tendency to form an alpha helix (h-region) between the n- and c-regions (Heijne 1990). This SP leads to the translocation of proteins by the conventional secretory pathway through the endoplasmic reticulum and the Golgi compartment. This pathway allows proteins to reach the extracellular space, be incorporated into the plasma membrane or targeted to an intracellular compartment such as the vacuole (Conesa *et al.* 2001; Shoji *et al.* 2008; Kubicek *et al.* 2014). Not

all secreted proteins contain a SP. Unconventional protein secretion pathways exist, translocating the proteins by either vesicles of an origin other than the Golgi compartment or direct translocation via transporter proteins (for more details, see the following reviews Nickel 2010; Ding *et al.* 2012; Girard *et al.* 2013; Rodrigues *et al.* 2013). The main function of extracellular proteins is to interact with the environment of a fungus, i.e., breaking down and acquiring nutrients, and to interact with other organisms, which is of particular importance for the interaction of a pathogen with its host (McCotter *et al.* 2016; Krijger *et al.* 2014).

The most important class of secreted proteins for the interaction of a pathogen with its host are effectors, which are virulence factors that are secreted into the cytoplasm of the host or the apoplast to suppress the host's immune response and manipulate its metabolism. In necrotrophic and hemibiotrophic fungi, effectors can also have toxic functions (Ciuffetti *et al.* 2010; Bent and Mackey 2007; Lo Presti *et al.* 2015; Sperschneider *et al.* 2015a). Most of the analysed effectors interact with the plant immune system as either a virulence (vir) factor or an avirulence (avr) factor. The plant immune response is initially triggered by a pathogen-associated molecular pattern (PAMP, for fungi usually chitin) on the cell surface. These PAMPs are recognised by membrane-localised receptors, leading to a defence response called PAMP-triggered immunity (PTI). For example, the apoplastic LysM domain containing effector Slp1 of *Magnaporthe oryzae*, is required for virulence by binding fungal chitin oligosaccharides; this binding suppresses the detection through plant receptors, which would lead to a chitin-induced PTI reaction (Mentlak *et al.* 2012). A similar mechanism was also shown for the Ecp6 effector of the ascomycete *Cladosporium fulvum* (Jonge *et al.* 2010). A mechanism other than the PTI, called effector-triggered immunity (ETI), involves the recognition of the secreted effectors by resistance-mediating receptor proteins (R-proteins), which trigger a strong defence response. In this case, the effector has an avirulence function. Most known effectors are characterized by their avr function; for example, the Avr2 protein of *C. fulvum* is secreted from the fungus to inhibit the tomato cysteine protease Rcr3. Such an interaction is detected by the plant's Cf receptor, which leads to an ETI reaction, in form of a hypersensitive response mediating resistance (Rooney *et al.* 2005; Wit, 2016). This co-evolution between the host and the pathogen leads to the formation of pathogenic races with differences in the effector repertory in the races. These differences are due to loss or modifications of effectors that can avoid the detection

through the corresponding arsenal of R-genes in the host plant. Meanwhile, the modification of the effector repertory is a selection pressure on the host for the formation of new R-gene variants (Bent and Mackey 2007).

This tight co-evolution between host and pathogen is a reason that makes the prediction of effectors in genomic sequences challenging. Because a pathogen adapts to one particular host, effectors rarely share sequence similarity with each other, but they do share some structural similarities. Effectors are often small, cysteine-rich, secreted proteins that are induced during infection (Lo Presti *et al.* 2015; Sperschneider *et al.* 2015a; Sperschneider *et al.* 2016).

Even though effectors often lack sequence similarity, a few motifs are shared by particular classes of effectors. An RxLR-motif, which is needed for translocation into the plant cell, is very commonly found after the SP in effectors of oomycetes and is often used for their identification in newly sequenced genomes (Whisson *et al.* 2007; Anderson *et al.* 2015). Unfortunately, no such highly conserved motif has been identified in fungi. Only the Y/F/WxC motif, previously found in the barley powdery mildew fungus, was also reported in effectors of rust fungi but exhibited less positional conservation (Duplessis *et al.* 2011; Godfrey *et al.* 2010).

In this study, we use the new sequenced draft genome of the *D. rosae* isolate DortE4 for an *in silico* prediction and analysis of its secretome with the goal of identifying possible key virulence factors. Based on this analysis, we will predict effector candidates and attempt to define a set of sequences that will serve as the starting point for further analysis, with the overall goal being to identify new virulence and avirulence factors that can aid in understanding the broad-spectrum resistance mediated by the Rdr1 locus and to identify new R-genes.

2. Materials and Methods

2.1 Sequence information

The gene models of the draft genome sequence (NCBI accession number: MVNX00000000) (Neu *et al.* 2017) of the isolate DortE4 were used for the prediction of the secretome and effector candidates.

2.2 Prediction of the secretome

The secretome was predicted in a manner similar to the guidelines of the Fungal Secretome KnowledgeBase (FunSecKB) (Meinken *et al.* 2016). For the prediction of the SP, the standalone version of SignalP 4.1 (Petersen *et al.* 2011) was used in combination with the Phobius web server (Kall *et al.* 2004). Sequences were predicted to contain a SP if both programs predicted the sequences to be secreted. To exclude membrane proteins, the TMHMM 2.0 web server (Krogh *et al.* 2001) was used in combination with transmembrane domain (TM) prediction by Phobius. Because the helical structure of the SP is often predicted to be a part of a TM, a protein was assumed to be secreted when a TM was predicted in the first 40 amino acids of the N-terminus. To exclude sequences that contain a SP but remain in the endoplasmic reticulum, all predicted secreted sequences were scanned for retention motifs from the PROSITE database (PS00014 ER_Targeting) with the ScanProsite web server (Castro *et al.* 2006). Web versions of WolfPSort (Horton *et al.* 2007) and TargetP 1.1 (Emanuelsson *et al.* 2007) were used to predict the subcellular localisation of sequences, but this information was not used to disqualify sequences. The PredGPI prediction server (Pierleoni *et al.* 2008) was used to predict secreted proteins that contain a GPI anchor.

2.3 Functional annotation of secretome sequences

The functional annotation of the predicted secretome was extracted from the whole genome annotation.

For a more detailed annotation of carbohydrate-degrading enzymes, the CAZy database (Lombard *et al.* 2014) and the dbCAN web server (Yin *et al.* 2012) were used. In addition, BLASTP (E-value cut-off e-10) was used to find matches in the pathogen-host interaction database (PHI-base) (Winnenburg *et al.* 2006; Winnenburg *et al.* 2008) for virulence factors and effectors.

2.4 Effector prediction pipeline

For the prediction of effector candidates three different strategies were combined. The basis for all analyses was the previously predicted secretome. CLC genomics workbench 9.0 (Qiagen, Hilden, Germany) was used to select sequences with a maximum length of 200 amino acids and a minimum cysteine content of 3%. The

stand-alone software EffectorP 1.0 (Sperschneider *et al.* 2016) was used in parallel to predict effector candidates. In a third approach, the web server of ScanProsite (Castro *et al.* 2006) was used to scan for the user-defined motif [YFW]-x-[C]. A sequence was included in the set of effector candidates when this motif was found within the first 90 amino acids of the protein.

2.5 Expression data

The analysed expression data were obtained from previously reported data (Neu *et al.* 2017). The raw reads are deposited in the Sequence Read Archive (SRA, SRX2494485-SRX2494496).

2.6 Analysis of duplicated effector candidates

BLASTP algorithm (Altschul *et al.* 1997) was used to identify paralogs of the predicted effector candidates in the whole proteome. To calculate identity scores between the candidates and their best matching sequences, multiple global sequence alignments were performed by means of the MAFFT 7.3 tool (Katoh and Standley 2013) using the FFT-NS-1 alignment algorithm and the software package trimAl 1.2 (Capella-Gutierrez *et al.* 2009).

3. Results and Discussion

3.1 Prediction of the whole secretome

The 14,004 predicted genes of the draft genome sequence of the *D. rosae* isolate DortE4 was the basis for the prediction of the secretome. A combination of different programs was used in this prediction (Figure 1). SignalP 4.1 (Petersen *et al.* 2011) and Phobius (Kall *et al.* 2004) were used for the prediction of SPs. Phobius, TMHMM 2.0 (Krogh *et al.* 2001) and ScanProsite (SP00014 ER_TARGET) were used to exclude membrane proteins and proteins that remain in the endoplasmic reticulum. Because TargetP uses the same algorithm for the prediction of SP as SignalP and WolfPSort has a low sensitivity (Sperschneider *et al.* 2015b), these data are only reported as additional information and not as disqualifiers.

With this pipeline, 827 proteins were predicted to be secreted (Figure 1, Supp. 1), which is 6% the whole predicted proteome of *D. rosae*. Zhu *et al*. (2012) predicted a similar fraction (5.97%) of secreted proteins in the genome of *Marssonina brunnea*, which is closely related to *D. rosae*, by using a combination of SignalP 3.0 and TMHMM 2.0. Krijger *et al.*(2014) used a combination of SignalP 2.0, 3.0, TargetP and TMHMM in a comparative study, resulting in a mean value of 7.6% for the fraction of proteins that are secreted from plant pathogenic fungi, which is slightly higher than the value found in *D. rosae.* However, the related fungi *Botrytis cinerea* and *Sclerotinia sclerotiorum* exhibited values of 5.6% and 5.5%, respectively, similar to the values of *D. rosae*. Different results were obtained in a broad analysis by Sperschneider *et al*. (2015b) of 48 fungal genomes, including nine from hemibiotrophic fungi with SignalP 3.0, TargetP and TMHMM. None of these nine fungi secreted fewer than 8% of their proteins.

Combining prediction programs increased specificity but decreased sensitivity, meaning that false positives were excluded and true positives were lost due to the stringency of the prediction (Min, 2010). Due to this fact, we report the results of WolfPSort (cut-off ≥ 17) (Horton *et al.* 2007) and TargetP (Emanuelsson *et al.* 2007) as additional information for predictions of greater stringency. These programs predicted that 615 respectively 817 of the 827 proteins are extracellularly localised (Figure 1, Supp. 1). Combining all programs resulted in predicting the secretion of 605 of the 827 proteins previously predicted to be secreted. The different strategies of the programs that were combined suggest that this set of sequences might be the most accurate but also the smallest.

PredGPI was used to predict GPI-anchored proteins and thus distinguish between these secreted proteins that become attached to the outer membrane and those that remain soluble. In total, 110 of the 827 secreted proteins were predicted to contain a GPI anchor, and the anchoring of 56 of them was classified as highly probable.

The computational prediction of secreted proteins should be seen as a method that can identify all proteins of this class. *In silico* predictions produce false positives as well as false negatives. Proteins secreted by unconventional protein secretion pathways can be translocated without a SP (Rodrigues *et al.* 2013). On the other hand, some proteins containing a SP can enter the secretory pathway but never reach the extracellular space because they are targeted to other compartments such

as the vacuole (Conesa *et al*. 2001). Experimental approaches such as proteomic analysis of the protein content of growth media or secretion trap assays are needed for the validation of these predictions and to identify proteins translocated by unconventional secretion pathways (Yang *et al.* 2012; Lee *et al.* 2012).

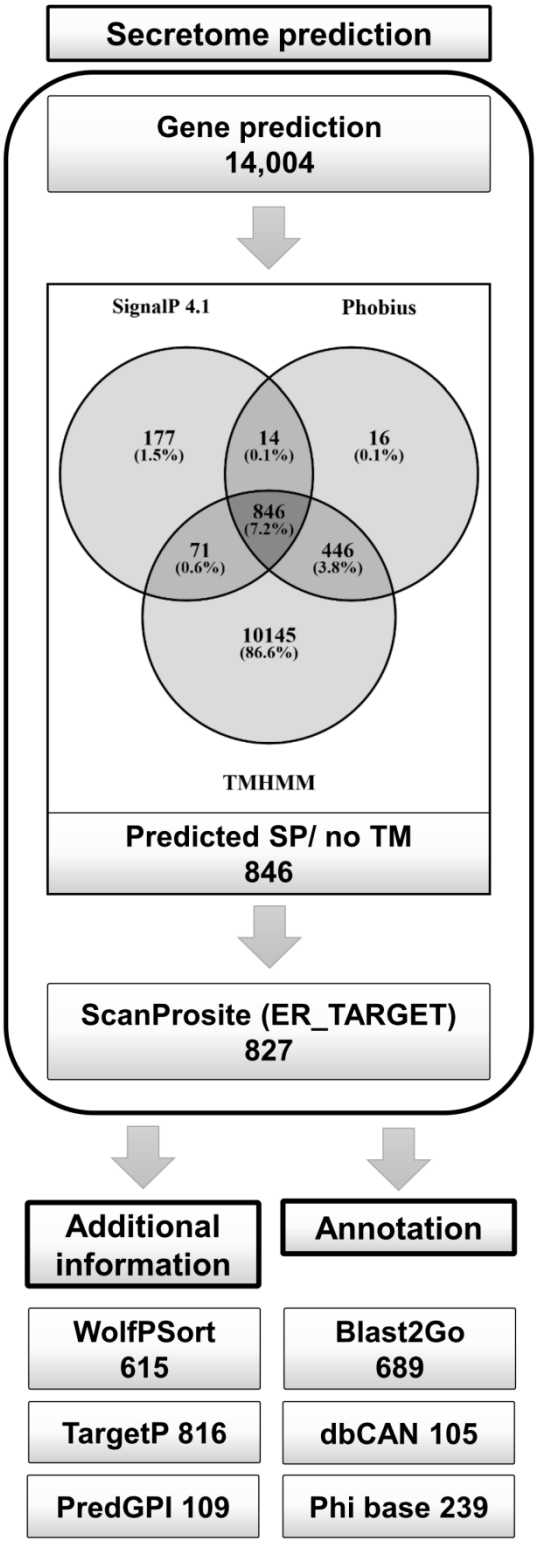


Figure 1: Secretome Prediction Pipeline. Proteins were classified as secreted when SignalP 4.1 and Phobius predicted a signal peptide (SP) and no transmembrane domain (TM) was predicted by Phobius or TMHMM to be past the first 40 amino acids of the protein. To exclude proteins that remain in the endoplasmic reticulum, we scanned the predicted proteins with ScanProsite for an ER-targeting signal. WolfPSort and TargetP 1.1 provided additional information for a more stringent prediction but

were not used to disqualify candidates. PredGPI was used to find proteins that are secreted but associated with the outer membrane by a GPI anchor. The functional annotation of the secretome was performed using Blast2GO and the dbCAN web server with the CAZy database for carbohydrate-active enzymes and the PHI-base.

3.2 Annotation of the secretome

The Blast2GO annotation (Gotz *et al.* 2008) assigned functional descriptions based on the top 20 BLAST matches for 696 of the 827 (84.16%) proteins of the secretome (Supp. 2), and 85 were annotated as hypothetical proteins, indicating that they are similar to other fungal proteins but that functional knowledge of them is lacking. The number of sequences with BLAST matches was comparatively small considering that 96.84% of the whole proteome was matched with other fungal proteins.

In addition to the functional descriptions assigned based on the BLAST matches, GO terms (Harris *et al.* 2004) were assigned to 566 sequences, and InterProScan results were assigned to 538 (Supp. 2). This additional information was used for enrichment analyses using Fisher's exact test implemented in the Blast2GO software to compare the composition of the secretome with that of the whole proteome. Both analyses showed that hydrolases are overrepresented in the secretome. For the GO-enrichment, 35.7% of the annotated sequences were assigned the GO term "hydrolases," and most of these sequences were annotated as involved in the hydrolysis of glycosyl bonds and proteolysis (Supp. 2). The secretome also contained sequences involved in the degradation of major components of the plant cell wall. All sequences of the proteome that were assigned the GO term "cutinase activity" were part of the secretome. More than half of the sequences with the GO term "pectate lyase activity" were predicted to be secreted. In addition, the GO terms "cellulose binding" and "cellulose catabolic process" were overrepresented, indicating that cellulases are part of the secretome. The enrichment also showed by the presence of GO terms such as "fungal-type cell wall" or "chitin binding" that the secretome contains proteins involved in the organization of the fungal cell wall.

This functional description of the *D. rosae* secretome shows strong similarities to the latest analysis of all secretomes registered at the FunSecKB (Meinken *et al.* 2016). A functional analysis of these data showed that major parts of the fungal secretome

contains hydrolases and that hydrolases involved in proteolysis and cell wall degradation are overrepresented, indicating that degradation of the plant cell wall is one of the major functions of the fungal secretome.

3.3 Expression of the secretome

Transcriptomic data of three early time points of the fungal infection (0, 24, 72 hpi) were used to generate an overview of the expressed portion of the secretome (Figure 2, Supp. 3). In total, 650 (78.7%) sequences of the secretome were expressed during at least one of the three time points, which is smaller than the percentage found for the whole proteome, where 88.5% of all sequences were expressed. Despite that difference, the expression of the secretome shows the same tendencies as the whole genome. Most sequences are expressed at 72 hpi, and almost one-third of the sequences were only detected with the RNAseq technique. In contrast to the sequences that were exclusively expressed at 72 hpi, most sequences were expressed either at 24 and 72 hpi or at all three time points. Very few sequences were exclusively detected at 0 or 24 hpi. As discussed in previous work, the reasons for these patterns might be both technical and biological. The amount of fungal biomass increases over time, as does the number of sequenced reads originating from the fungus. Many genes might thus not be detected at the early time points, even if they are expressed. However, drastic changes that substantially affect the transcriptome occur in the biology of the fungus during these stages of its development. During the first 24 hours, the spores are germinating and develop appressoria to penetrate the plant cell. Afterward, haustoria are formed to acquire nutrients and interact with the host cell. The fungus is in a phase of rapid growth during the development of these specific structures. All these processes are accompanied by transcriptional changes. A transcriptomic study of a *S. sclerotiorum* infection of *Brassica napus* demonstrated that between 2.6 and 8.8% of all expressed fungal genes were up-regulated in the first 48 hours (Seifbarghi *et al.* 2017).

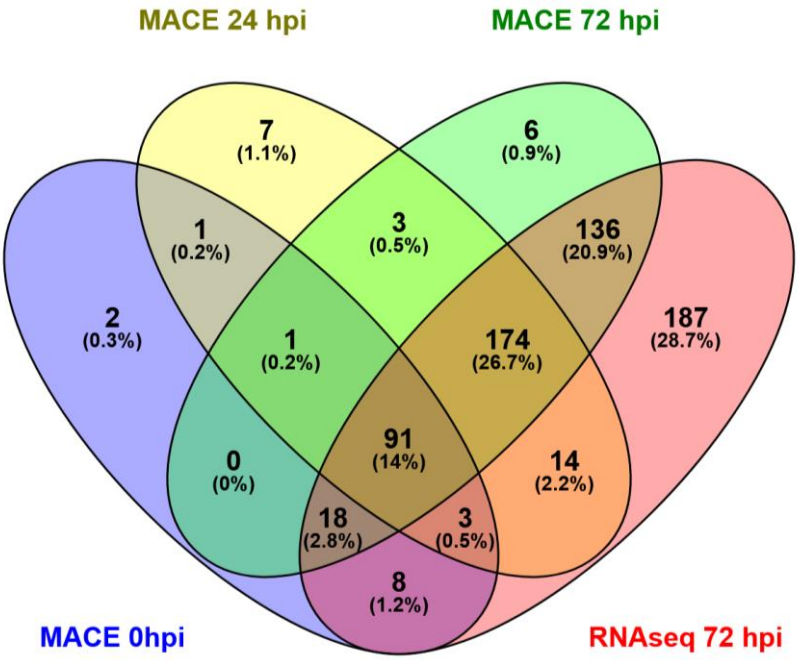


Figure 2: Venn diagram of the expression of the secretome sequences during the early stages of infection (0, 24, and 72 hpi). Data were generated by means of the massive analysis of cDNA ends (MACE) and conventional RNAseq.

3.4 Potential virulence factors

One of the crucial steps during the development of a fungal pathogen is its entry into the plant cell. Many enzymes involved in the breakdown of the plant cell wall are thus virulence factors. To better understand the capacity of the secretome to degrade the plant cell wall, we combined different sources of functional information such as Blast2GO functional descriptions, the assigned EC-numbers, GO terms and InterPro IDs as well as the results of an annotation with the CAZy database, which was used to find plant cell wall-degrading enzymes (CWDEs) (Supp. 4). The results are summarised in Table 1. The secretome contains 111 enzymes that collectively target the degradation of all the main components of the plant cell wall as well as 17 lipases, indicating that the secretome plays a crucial role in decomposing the plant cell wall and the consequent penetration process. Most of these enzymes are involved in the degradation of hemicellulose. Hemicellulose is a mixture of polysaccharides containing different sugars, including xylose, arabinose, galactose, mannose and rhamnose. Different enzymes are thus needed to break down this cell

wall component. In addition to this mixture of enzymes, most CWDEs are involved in the degradation of cellulose and pectin, which are the main components of the primary cell wall and the middle lamella. The secretome also contains 12 cutinases and only four callose- and lignin-degrading enzymes.

Members of all classes of CWDEs are expressed during the early stages of plant entry (Supp. 3 and 4). Almost all potential hemicellulose-, cellulose- and pectin-degrading enzymes are expressed during at least one of the analysed time points. Some of the pectin-degrading enzymes derive from the genes with the highest levels of expression at 24 and 72 hpi, indicating that these genes might be of special importance for this process. Only 6 of 12 cutin-degrading enzymes and one of the four lignin-degrading enzymes were generally expressed only at 72 hpi with comparable low expression values.

Brown *et al*. (2012) analysed the secretome of the hemibiotrophic pathogen *Fusarium graminearum* grown on different cereals in a manner similar to that presented here. The amounts of CWDEs affecting the different cell wall components were similar to those in the present data except that fewer cutinases and pectin/pectate-degrading enzymes were present in *F. graminearum* than the black spot secretome, which is not surprising because lower levels of pectin-degrading enzymes are often reported for pathogens of monocots due to differences in cell wall composition (Kubicek *et al.* 2014).

Table 1: Overview of the number of CWDEs in the secretome.

Target of degradation	Number of secreted proteins
Cutin	12
Cellulose/Cellubiose	23
Hemicellulose	47
Callose	4
Lignin	4
Pectin/Pectate	21
Lipids/Choline	17
Total	**128**

A BLASTP (E-Value 1E-10) with the PHI-base (Winnenburg *et al.* 2006; Winnenburg *et al.* 2008) was performed to more closely examine the pathogenic features of the secretome (Supp. 4). In total, 239 proteins of the predicted secretome matched with sequences from the PHI-base, and 81 of these matches are CWDEs. Accessions in the database are classified into different categories due to the results of mutation experiments. The categories indicating virulence factors, which are the most relevant, are “loss of pathogenicity and reduced virulence”. Only 18 sequences of the black spot secretome matched proteins classified with “loss of pathogenicity,” and 115 of the sequences were associated with the class “reduced virulence”.

Six of the 18 sequences with the classification “loss of pathogenicity” are involved in the degradation of the cuticle, and most of them were expressed. These sequences might be important virulence factors, similar to the cutinases of other plant pathogenic fungi (Eleonora Kikot *et al.* 2009; Skamnioti and Gurr 2007; Wang *et al.* 2017). In addition, two autophagy lipases matched a protein from *M. oryzae* (PHI:2081), and two EF-domain-containing proteins matched an EF-domain-

containing protein from *S. sclerotiorum* (PHI:3936); both are required for appressorium formation and are crucial for entry of the pathogens into plants (Kershaw and Talbot 2009; Xiao *et al.* 2014). These similarities hint that orthologs in *D. rosae* may also be involved in appressorium formation. The expression data support this hypothesis because the expression of three of these four sequences was detected at 24 hpi where the first appressoria had already formed, but the expression values were too low to substantively indicate induction at this stage. Nevertheless, the described sequences are interesting starting points for the functional analysis of potential virulence factors.

In addition, the sequences of enzymes involved in the degradation of pectin, callose, hemicellulose and lignin were homologous to sequences in the category “reduced virulence” from the PHI-base. In addition to plant cell wall degradation, putative virulence factors are involved in various other processes such as fungal cell wall organisation, proteolysis and stress responses.

3.5 Prediction of effectors

The prediction of fungal effectors is a challenging task since effectors of different species share very little sequence similarity, which is due to the co-evolution of a pathogen and its host. This reason is why prediction approaches normally use the structural characteristics of a protein or a conserved sequence motif as an indicator for potential effector proteins. However, the prior definition of these characteristics is not consistent within the scientific community, and not all effector proteins share all of these characteristics. For this reason, we combined three different approaches for our prediction pipeline. Figure 3 gives an overview of the prediction pipeline and its results. Because many known effectors belong to the class of small secreted cysteine-rich proteins (SSP), we used the characteristics of this class as indicators for effector proteins.

The definition of these characteristics differs depending on the author, sometimes quite significantly. Saunders *et al*. (2012), e.g., used 150 amino acids as a cut-off for small proteins and defined those with more than 3% cysteine as enriched. In contrast, other authors used different cut-offs between 200 and 400 amino acids, (Rep 2005; Bowen *et al.* 2009; Hacquard *et al.* 2012) to define a protein as small.

Here, we used a definition similar to those used by Saunders *et al*. (2012), specifically a maximal protein size of 200 amino acids and a minimum of 3% cysteine content in the protein. In total, 114 secreted proteins fulfilled these criteria. However, some known effectors do not fulfil these criteria (Sperschneider *et al.* 2015a). To include other criteria, we used the prediction program EffectorP (Sperschneider *et al.* 2016), which uses a machine learning approach to build a model for a prediction based on a variety of features that discriminate a set of known effector proteins from other secreted proteins. This pipeline resulted in the prediction of 171 potential effector proteins. We also used an approach based on the sequence motif Y/F/WxC, which was first discovered in the barley powdery mildew fungus *Blumeria graminis* (Godfrey *et al.* 2010) and later in a less-conserved position in rust fungi (Duplessis *et al.* 2011; Saunders *et al.* 2012). We identified 122 proteins carrying this motif in their N-terminal regions in the secretome.

With the described pipeline, we predicted 244 effector candidates in the secretome of *D. rosae*, while 52 were predicted by all three approaches (Figure 3, Supp. 5), making these 52 predictions the most reliable and the best subjects for initial detailed analyses. This set is of particular interest, because of the occurrence of the Y/F/WxC motif, indicating that there is a group of effectors shared by fungi of different lifestyles. So far the motif is only reported for obligate biotrophic fungi and *D. rosae* has a hemibiotrophic lifestyle. So it might be, that these effectors function in its biographic stage, where it forms haustoria. The results of EffectorP and the utilisation of characteristics of a SSP greatly overlap, which might be due to the fact that several features EffectorP uses for model building are closely related to the criteria defined for SSPs. The most discriminative features used by EffectorP are the molecular weight, sequence length, protein net charge and percentages of cysteines, serines and tryptophans in the sequence (Sperschneider *et al.* 2016). In addition, the majority of the proteins that EffectorP uses as a training dataset were identified by using the characteristics of SSPs. It is therefore not surprising that almost all potential SSPs were included in the EffectorP prediction results. However, the program predicted 72 additional effector candidates. Almost the same number of additional candidates were identified by searching for the Y/F/WxC-motif in the N-terminal regions of proteins, but only a small fraction of these candidates were also predicted with one of the two other approaches. Only 12 were also predicted by EffectorP or shared characteristics of SSPs.

In addition to the sequences identified by these prediction approaches, seven sequences from the secretome matched with known effectors in the PHI-base (Supp. 4), but none of them were predicted using any of the applied prediction approaches. Nevertheless, a BLAST match is a valuable hint for the function of a protein, so these sequences were included in the set of effector candidates. With these seven additional candidates, the total number of effector candidates was 251.

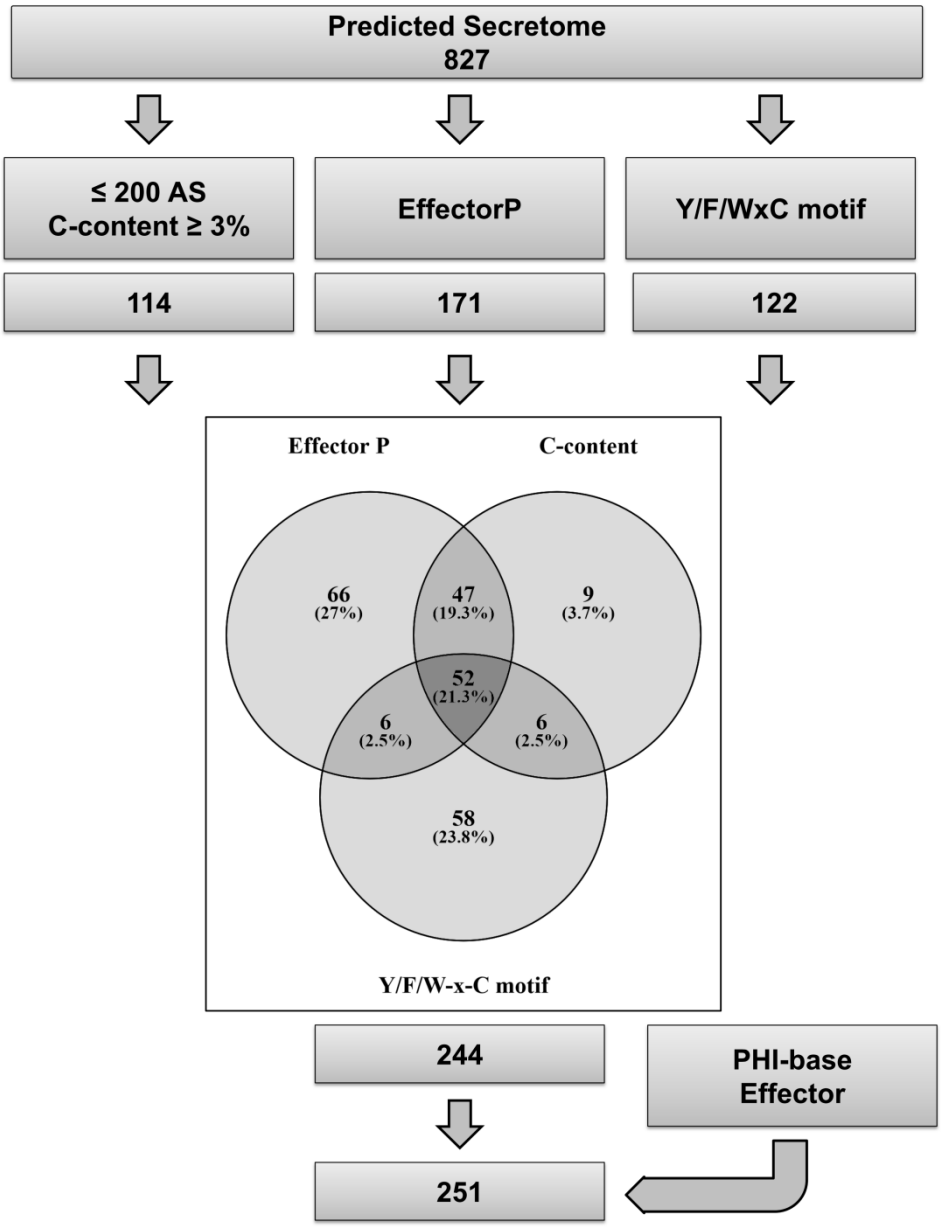


Figure 3: Prediction pipeline for effector proteins of *D. rosae*. Three approaches were combined for the prediction of effector candidates: characteristics of SSPs, prediction with EffectorP software and the presence of a Y/F/WxC motif at the N-terminal region of a protein. In addition to these approaches, a BLASTX against sequences of known effectors from the PHI-base was performed.

3.6 Annotation of the effector candidates

The BLAST annotation does not result in any matches for 108 of 251 effector candidates, indicating that these sequences are specific for the black spot genome. This fact is not surprising because the evolution of effectors is tightly linked to the evolution of the host immune response. Every pathogen thus needs a specific set of effectors that are highly specialized for interactions with its host plant. The absence of BLAST hits is sometimes taken as a prediction parameter for possible effector proteins (Syme *et al.* 2013; Bowen *et al.* 2009), indicating that these unannotated sequences might be the most interesting sequences for further analysis. In addition to the candidates without BLAST matches, 40 sequences match those of proteins of other fungi that were not functionally annotated ("hypothetical protein") or had annotations with little functional information such as "small secreted protein" "extracellular serine-rich protein" or "signal peptide-containing protein". The majority of the 52 candidates predicted by all three prediction approaches were not matched to other fungal proteins by BLAST annotation. As mentioned before, this result is a hint that these proteins are highly specific for *D. rosae*, supporting the value of this dataset as a promising starting point for further analyses.

Nevertheless, 108 sequences are functionally annotated. Many of them are involved in cell wall degradation or cell adhesion, indicating that these proteins might be virulence factors but not effectors in the sense that they are involved in interactions with the plant immune response. The three candidates containing a chitin-binding domain might be of particular interest because the interactions of effectors with chitin to suppress the chitin-induced defence response is a mechanism used by known effector proteins such as Avr4 or Ecp6 (van den Burg *et al.* 2006; Jonge *et al.* 2010). Surprisingly, none of these chitin-binding proteins contains a LysM domain, which is found in the Ecp6 effector of *C. fulvum* and effector candidates of the related species *M. brunnea* (Jiang *et al.* 2014). Additional candidates were found by a BLAST against the PHI- Four of them were annotated as serine endopeptidases (DR_00434, DR_012785, DR_004799, and DR_006904), which is consistent with their PHI-base matches, GIP1 (PHI:652) and GIP2 (PHI:653) of the oomycete *Phytophthora sojae* (Rose *et al.* 2002). These proteins are inactive serine endopeptidases acting as inhibitors of the endo-beta-1,3-glucanases of the host to suppress its elicitor-mediated defence response, which is normally triggered by glucan oligosaccharides

from the fungal cell wall. The corresponding sequences in the black spot secretome might function similarly in the interaction of the fungus with its host. Two other effector candidates shared similarity with an effector of the PHI-base. DR_008872 and DR_010809 were both annotated as ricin b lectins and were matched by BLAST to the apoplastic effector MoCDIP1 (PHI:3213) of *M. oryzae*, which induces cell death (Chen *et al.* 2013). Because both pathogens have a hemibiotrophic lifestyle, such a cell death-inducing gene might have an important role during the necrotrophic stage of the pathogens.

3.7 Expression of effector candidates

Effector genes are often induced during the interaction with plants. The presented expression data are a useful tool to differentiate which candidates are the most promising for further analysis, even if the accurate quantification of very low expression levels is problematic due to the technical issues mentioned above.

In total, 204 of the 251 effector candidates were expressed during at least one of the three time points (0, 24, and 72 hpi) in the early stages of infection. Only one of the 52 effector candidates that had been predicted with all three approaches was not detected in the transcriptomic data. The data show that sequences without any functional annotation are represented particularly strongly in the most expressed sequences in the set of candidates at 24 and 72 hpi, indicating that they might be induced at the time of penetration and haustoria formation.

This effect becomes even more interesting when the expression data of the whole proteome are considered; 15 effector candidates are among the 100 strongest expressed sequences at 24 hpi, and 6 among those at 72 hpi. All of these highly expressed candidates lack any functional annotation, and eight of them belong to the 52 sequences predicted by all approaches (Supp. 5). Five of these highly expressed sequences (DR_002828, DR_003618, DR_009552, DR_006285, and DR_003215) are remarkable due to their expression pattern. DR_002828 is the strongest induced effector at 24 hpi and has the fourth highest expression level of the whole proteome at this time point. Its expression value drastically drops at 72 hpi, indicating that it might be more important at the beginning of the interaction. At 72 hpi, the effector candidates DR_006285, DR_003215, and DR_003618 are the strongest induced

genes in the whole *D. rosae* proteome. All of them are also among the 100 most expressed genes at 24 hpi. This expression level might indicate an induction of these genes in haustoria, which start their development at 24 hpi. Haustoria become larger, and their number increases over time with the development of more and more mycelia. DR_00952 was also expressed at very high levels at both time points. Even if its expression level was slightly lower than that of the other three candidates, it can be still considered to be induced in haustoria.

The expression patterns of the annotated effector candidates are less remarkable. Only the two ricin-b-like lectins, which match the effector MoCDIP1 in the PHI-base, had high expression levels at 24 and 72 hpi indicating an induction. Only one of the four Ser-peptidases is expressed at 24 hpi. However, at 72 hpi, all four were detected, and one of them belongs to the 10% of genes with the highest expression level at this time point.

3.8 Influence of genome duplication

In previous analyses of the *D. rosae* genome, we found that a large portion of the genome had been duplicated (Neu *et al.* 2017). Because gene duplication and diversification is a typical mechanism in the evolution of effectors (Stergiopoulos *et al.* 2012; Lo Presti *et al.* 2015; Selin *et al.* 2016), we analysed the degree of gene duplication within the set of effector candidates.

A BLASTP search of all effector candidates in the whole proteome resulted in 223 (88.8%) sequences with at least one match (Supp. 6). Multiple global sequence alignments allowed us to calculate the sequence identities shared by the effector candidates and their potential paralogs. Only 49% of the predicted proteins and 46% of their corresponding mRNA sequences share more than 70% identity with their most similar match. This number is surprisingly low compared to the number of duplicated benchmarking universal single-copy orthologs (BUSCOs) that were used for the analysis of genome duplication. Here, more than 84% of the duplicates shared at least 70% identity. This result indicates that diversity within duplicated effector genes is much higher than that in other classes of proteins, which is a hint that diversification is already in progress. The process of duplication and diversification is also shown to be one of the mechanisms forming the effector composition of a

pathogen by the fact that 65% of the effector candidates share the most identity with other candidates and that additional10.7% are best matched with another protein of the secretome that might be an effector but does not match the prior defined criteria. Some of the candidates that are of particular interest due to their induction during penetration and haustorium development or their similarity to known effectors also occur in pairs. DR_006285 and DR_003215 share approximately 70% identity. The sequences of DR_008872 and DR_010809 match that of the MoCDIP1 effector, DR_12785 and DR_004799 match the GIP1-effector, and DR_006904 and DR_004344 match the GIP2-effector; each of these gene pairs share identity values within the pair of 85.9% to 97.7%.

On the other hand, DR_009552 and the other 27 candidates that do not match any other sequences in the proteome, even with BLAST, are quite interesting because this lack of matching might indicate a loss of one of the duplicates, which we have already observed for some of the analysed BUSCOs. This loss differed between different isolates of *D. rosae* (Neu *et al.* 2017). A comparison of duplicated effector candidates in different isolates can provide more insight into the evolution of these effectors and, if duplication is involved, into the formation of pathogenic races.

4. Conclusion

The secretome is the main interface for the interaction of a pathogen with its host and includes the most important virulence factors and effectors. We combined different approaches and used the newly published draft genome sequences of *D. rosae* isolate DortE4 to predict the secretion of 827 proteins, including many enzymes involved in the degradation of all major components of the plant cell wall. Some pectin-degrading enzymes were noticeable due to their high expression levels.

The most important class of secreted proteins are the effector proteins, which are secreted from a pathogen into the apoplast or the cytoplasm of the host cell to influence the defence response of the plant. We identified 251 effector protein candidates. A subset of 52 candidates is of particular interest because these proteins were predicted by EffectorP software, share all the characteristics of SSPs and contain a Y/F/WxC motif in their N-terminal region. Especially the occurrence of this

sequence motif is astonishing, because so far it has only been reported in obligate biotrophic fungi (Duplessis *et al.* 2011; Godfrey *et al.* 2010), indicating that this class of effector proteins is shared by biotrophic and hemibiotrophic fungi. Further analysis of this set of effector candidates could validate whether this motif can be used as an identifier in fungi, similar to the RxLR-motif of oomycetes.

Furthermore, we pointed out 11 additional candidates, five of whom showed an outstandingly strong transcriptional induction during penetration and haustorium formation and six whom share sequence similarity with known effector proteins of other fungi. Extensive analysis of roses with the sequenced black spot isolate DortE4 revealed resistant genotypes with R-genes other than Rdr1 (unpublished data). The presented set of effector candidates can now be used to screen for new R-genes and to find the avirulence factors recognised by the Rdr1-gene and other R-genes by techniques such as knock-outs, transient expression, interaction analysis and comparisons of the effector content of different *D. rosae* isolates (approaches reviewed in Dalio *et al.* 2018).

Acknowledgment

We thank Jasper Rees and Jonathan Featherston for their cooperation on the genome sequence of *D. rosae* and their continuous interest in this project. We are grateful for the funding by the “Deutsche Forschungsgemeinschaft” (DFG) through the GRK1798 “Signaling at the Plant-Soil Interface”. The authors state that there are no conflicts of interest related to this publication.

Supporting Information:

Supp. 1: Results of the secretome prediction pipeline

Supp. 2: Annotation of the *D. rosae* secretome

Supp. 3: Expression data of the secretome

Supp. 4: Plant CWDEs and PHI-base data

Supp. 5: Effector prediction pipeline

Supp. 6: Analysis of effector paralogs

References

Altschul S.F., Madden T.L., Schaffer A.A., Zhang J.H., Zhang Z., Miller W., Lipman D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *NAR* , 25(17), 3389–3402.

Anderson R.G., Deb D., Fedkenheuer K., McDowell, J.M., 2015. Recent Progress in RXLR Effector Research. *MPMI*, **28(**10), 1063–1072.

Aronescu A., 1934. Diplocarpon rosae: from spore germination to haustorium formation, Diss. New York Columbia Univ.

Bent A.F., Mackey D., 2007. Elicitors, effectors, and R genes: the new paradigm and a lifetime supply of questions. *Annu. Rev. Phytopathol.,* 45, 399–436.

Blechert O., Debener, T. 2005. Morphological characterization of the interaction between Diplocarpon rosae and various rose species. *Plant Pathol.*, 54(1), 82–90.

Bowen J.K., Mesarich C.H., Rees-George J., Cui W., Fitzgerald A., Win J., Plummer K.M.,Templeton, M.D., 2009: Candidate effector gene identification in the ascomycete fungal phytopathogen Venturia inaequalis by expressed sequence tag analysis. *Mol. Plant. Pathol.*, 10(3). 431–448.

Brown N.A., Antoniw J., Hammond-Kosack, K.E., 2012. The Predicted Secretome of the Plant Pathogenic Fungus Fusarium graminearum: A Refined Comparative Analysis. *PloS one*, 7**(**4), e33731

Capella-Gutierrez S., Silla-Martinez J.M., Gabaldon T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinf.*, 25(15), 1972–1973.

Castro E. de, Sigrist C.J.A., Gattiker A., Bulliard V., Langendijk-Genevaux P.S., Gasteiger E., Bairoch A. and Hulo, N., 2006. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *NAR*, 34, W362-365.

Chen S., Songkumarn P., Venu R.C., Gowda, M., Bellizzi M., Hu J., Liu W., Ebbole D., Meyers B., Mitchell T., Wang, G.-L., 2013. Identification and characterization of in planta-expressed secreted effector proteins from Magnaporthe oryzae that induce cell death in rice. *MPMI*, 26(2), 191–202.

Ciuffetti L.M., Manning V.A., Pandelova I., Betts M.F., Martinez, J.P., 2010. Host-selective toxins, Ptr ToxA and Ptr ToxB, as necrotrophic effectors in the Pyrenophora tritici-repentis-wheat interaction. *New Phyto.*, 187(4), 911–919.

Conesa A., Punt P.J., van Luijk N., van den Hondel, C.A. 2001. The secretion pathway in filamentous fungi: a biotechnological view. *Fungal Genetics and Bio.*, 33(3), 155–171.

Dalio R. J. D., Herlihy J., Oliveira T.S., McDowell J.M., Machado, M., 2018. Effector Biology in Focus: A Primer for Computational Prediction and Functional Characterization. *MPMI*, 31**(**1), 22–33.

Debener, T., Byrne D.H., 2014. Disease resistance breeding in rose: current status and potential of biotechnological tools. *Plant Science*, 228, 107–117.

Debener, T., Drewes-Alvarez, R., Rockstroh, K. 1998. Identification of five physiological races of blackspot, Diplocarpon rosae, Wolf on roses. *Plant Breeding*, 117**(**3) 267–270.

de Wit P.J.G.M.,2016. Cladosporium fulvum Effectors: Weapons in the Arms Race with Tomato. *Annu. Rev. Phytopathol*., 54, 1–23.

Ding Y., Wang J., Wang J., Stierhof Y.-D., Robinson D.G., Jiang L., 2012. Unconventional protein secretion. *Trends in Plant Sci.*, **17**(10), 606–615.

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., Kues 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., Rouze P., Ellis J.G., Dodds P.N., Schein J.E., Zhong S., Hamelin R.C., Grigoriev I.V., Szabo L.J., Martin F., 2011. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proc. Natl. Acad. Sci.*, 108**(**22), 9166–9171.

Kikot E. G., Hours A. R., Alconada M. T., 2009. Contribution of cell wall degrading enzymes to pathogenesis of Fusarium graminearum: a review. *J. of Basic Microbio*., 49(3) 231–241.

Emanuelsson O., Brunak S., Heijne G. von, Nielsen, H., 2007. Locating proteins in the cell using TargetP, SignalP and related tools. *Nature Prot*., 2(4), 953–971.

Frick L. 1943. Untersuchungen über Biologie und Pathogenität von *Diplocarpon rosae* (Lib.) *Wolf*. Diss. ETH Zurich

Gachomo E.W., Dehne H. W., Steiner U., 2006. Microscopic evidence for the hemibiotrophic nature of Diplocarpon rosae, cause of black spot disease of rose. *PMPP*.,69(1), -386–92.

Gachomo E.W., Kotchoni, S.O., 2007. Detailed description of developmental growth stages of Diplocarpon rosae in Rosa: a core building block for efficient disease management. *Annals of App. Bio.*, (151), 2233–243.

Gachomo E.W., Seufferheld M.J., Kotchoni, S.O. 2010. Melanization of appressoria is critical for the pathogenicity of Diplocarpon rosae. *Mol. Bio. Rep.,* 37**(**7), 3583–3591.

Girard V., Dieryckx C., Job C., Job D., 2013. Secretomes: the fungal strike force. *Proteomics*, 13**(**3-4), 597–608.

Godfrey, D., Bohlenius, H., Pedersen, C., Zhang, Z., Emmersen, J. and Thordal-Christensen, H. (2010) Powdery mildew fungal effector candidates share N-terminal Y/F/WxC-motif. *BMC Genomics*, **11**, 1 317.

Gotz S., Garcia-Gomez J.M., Terol,J., Williams T.D., Nagaraj S.H., Nueda M.J., Robles M., Talon M., Dopazo J., Conesa A., 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *NAR*, 36**(**10) 3420–3435.

Hacquard S., Joly D.L., Lin Y.-C., Tisserant E., Feau N., Delaruelle C., Legue V., Kohler A., Tanguay P., Petre B., Frey P., van de Peer Y., Rouze P., Martin F., Hamelin R.C., Duplessis S., 2012. A Comprehensive Analysis of Genes Encoding Small Secreted Proteins Identifies Candidate Effectors in Melampsora larici-populina (Poplar Leaf Rust). *MPMI*, 25, 3279–293.

Harris M.A., Clark J., Ireland A., Lomax J., Ashburner M., Foulger R., Eilbeck K., Lewis S., Marshall B., Mungall C., Richter J., Rubin G.M., Blake J.A., Bult C., Dolan M., Drabkin H., Eppig J.T., Hill D.P., Ni L., Ringwald M., Balakrishnan R., Cherry J.M., Christie K.R., Costanzo M.C., Dwight S.S., Engel S., Fisk D.G., Hirschman J.E., Hong E.L., Nash R.S., Sethuraman A., Theesfeld C.L., Botstein D., Dolinski K., Feierbach B., Berardini T., Mundodi S., Rhee S.Y., Apweiler R., Barrell D., Camon E., Dimmer E., Lee V., Chisholm R., Gaudet P., Kibbe W., Kishore R., Schwarz E.M., Sternberg P., Gwinn M., Hannick L., Wortman J., Berriman M., Wood V., La Cruz N. de, Tonellato P., Jaiswal P., Seigfried T., White, R., 2004. The Gene Ontology (GO) database and informatics resource. *NAR*, 32, D258-261.

Hattendorf A., Linde M., Mattiesch L., Debener T., Kaufmann H., 2004. Genetic Analysis of Rose Resistance Genes and their Localisation in the Rose Genome. *Act. Hort*., 651, 123–130.

Heijne G., 1990. The signal peptide. *J. of Membrane Bio*., 115(3), 195–201.

Horst, R.K., Cloyd, R.A.,2007. Compendium of rose diseases and pests. 2nd ed. St. Paul, MN: APS Press.

Horton P., Park K.-J., Obayashi,T., Fujita, N., Harada H., Adams-Collier C.J., Nakai K., 2007. WoLF PSORT: protein localization predictor. *NAR*, 35, W585-587.

Jiang C., He B., Huang,R., Huang M., Xu L., 2014. Identification and functional analysis of LysM effectors from Marssonina brunnea. *Australasian Plant Pathol*., 43, 6615–622.

Jonge R. de, van Esse H.P., Kombrink A., Shinya T., Desaki Y., Bours R., van der Krol S., Shibuya N., Joosten M.H.A.J., Thomma, B.P.H.J., 2010. Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. *Science*, 329, 5994 953–955.

Kall L., Krogh A., Sonnhammer E.L.L., 2004. A combined transmembrane topology and signal peptide prediction method*. J. of Mol. Bio*., 338, 51027–1036.

Katoh K.,Standley D.M., 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *MBE* , 30**(**4), 772–780.

Kaufmann H., Mattiesch L., Lorz H., Debener, T., 2003. Construction of a BAC library of Rosa rugosa Thunb. and assembly of a contig spanning Rdr1, a gene that confers resistance to blackspot. *Mol. Genetics and Genomics*, 268(5), 666–674.

Kershaw, M.J. and Talbot, N.J. (2009) Genome-wide functional analysis reveals that infection-associated fungal autophagy is necessary for rice blast disease. *Proc. Natl. Acad. Sci.*, 106, 37 15967–15972.

Krijger J.J., Thon M.R., Deising H.B., Wirsel S.G.R., 2014. Compositions of fungal secretomes indicate a greater impact of phylogenetic history than lifestyle adaptation. *BMC Genomics*, 15(1), 722.

Krogh, A., Larsson B., von Heijne G., Sonnhammer E.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. of Mol. Bio.*, 305(3), 567–580.

Kubicek C.P., Starr T.L., Glass, N.L., 2014. Plant Cell Wall-Degrading Enzymes and Their Secretion in Plant-Pathogenic Fungi. *Annu. Rev. Phytopathol*., **(**52), 427–451.

Lee S.J., Rose, J., 2012. A yeast secretion trap assay for identification of secreted proteins from eukaryotic phytopathogens and their plant hosts. *Meth. in Mol. Bio*., 835**,** 519–530.

Lo Presti L., Lanver D., Schweizer G., Tanaka S., Liang L., Tollot M., Zuccaro A., Reissmann S., Kahmann R., 2015. Fungal effectors and plant susceptibility. *Annu. Rev. Plant Bio*., 66, 513–545.

Lombard V., Golaconda Ramulu H., Drula E., Coutinho P.M., Henrissat B., 2014.The carbohydrate-active enzymes database (CAZy). *NAR*, 42, D490-D495.

McCotter S.W., Horianopoulos L.C., Kronstad J.W.,2016. Regulation of the fungal secretome. *Cur.Genetics*, 62**(**3), 533–545.

Meinken J., Asch D.K., Neizer-Ashun K.A., Chang G.H., Cooper JR. C., R., Min, X.J., 2016. FunSecKB2: A fungal protein subcellular location knowledgebase. *Computational Mol. Bio*., 4(4)

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., Talbot N.J., 2012. Effector-mediated suppression of chitin-triggered immunity by magnaporthe oryzae is necessary for rice blast disease. *Plant Cell*, 24(1) 322–335.

Menz I., Straube J., Linde M., Debener T., 2017. The TNL gene Rdr1 confers broad-spectrum resistance to Diplocarpon rosae. *Mol. Plant Pathol.*, 19(5), 1104-1113

Min, X.J., 2010. Evaluation of Computational Methods for Secreted Protein Prediction in Different Eukaryotes. *JPB*, 3, 143–147.

Münnekhoff A.K., Linde M., Debener, T., 2017. The gene diversity pattern of Diplocarpon rosae populations is shaped by the age, diversity and fungicide treatment of their host populations. *Plant Pathol.*, 66(8), 1288-1298.

Neu E., Featherston J., Rees J., Debener, T., 2017. A draft genome sequence of the rose black spot fungus Diplocarpon rosae reveals a high degree of genome duplication. *PloS one*, 12(10), e0185310.

Nickel W., 2010. Pathways of unconventional protein secretion. *Cur. Op. in Biotech*, 21(5), 621–626.

Petersen T.N., Brunak S., von Heijne G., Nielsen H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Meth.*, 8(10), 785–786.

Pierleoni A., Martelli P.L., Casadio R., 2008. PredGPI: a GPI-anchor predictor, *BMC Bioinfo*., 9(1), 392.

Rep M., 2005. Small proteins of plant-pathogenic fungi secreted during host colonization. *FEMS Microbio. Letters*. 253(1), 19–27.

Rodrigues M.L., Franzen A.J., Nimrichter L., Miranda, K., 2013. Vesicular mechanisms of traffic of fungal molecules to the extracellular space. *Cur. Op. in Microbio*., 16(4), 414–420.

Rooney H.C.E., Van't Klooster J.W., van der Hoorn R.A.L., Joosten M.H.A.J., Jones J.D.G., de Wit, P.J.G.M., 2005. Cladosporium Avr2 inhibits tomato Rcr3 protease required for Cf-2-dependent disease resistance. *Science*, 308(5729), 1783–1786.

Rose J.K., Ham K.S., Darvill A.G., Albersheim P.,2002. Molecular cloning and characterization of glucanase inhibitor proteins: Coevolution of a counterdefense mechanism by plant pathogens. *Plant Cell*, 14(6), 1329–1345.

Saunders D.G.O., Win J., Cano L.M., Szabo L.J., Kamoun S., Raffaele, S., 2012. Using Hierarchical Clustering of Secreted Protein Families to Classify and Rank Candidate Effectors of Rust Fungi. *PloS one*, 7(1), e29847.

Seifbarghi S., Borhan M.H., Wei Y., Coutu C., Robinson S.J., Hegedus, D.D., 2017. Changes in the Sclerotinia sclerotiorum transcriptome during infection of Brassica napus. *BMC Genomics*, 18(1), 266.

Selin C., de Kievit T.R., Belmonte M.F., Fernando, W. G. Dilantha, 2016. Elucidating the Role of Effectors in Plant-Fungal Interactions: Progress and Challenges. *Front. Microbio.*, 7, 600.

Shoji J.Y., Arioka M., Kitamoto, K., 2008. Dissecting cellular components of the secretory pathway in filamentous fungi: insights into their application for protein production, *Biotech. Letters*, 30(1), 7–14.

Skamnioti P., Gurr S.J., 2007. Magnaporthe grisea cutinase2 mediates appressorium differentiation and host penetration and is required for full virulence. *Plant Cell*, 19(8) 2674–2689.

Sperschneider J., Dodds P.N., Gardiner D.M., Manners J.M., Singh K.B., Taylor J.M., 2015a. Advances and challenges in computational prediction of effectors from plant pathogenic fungi. *PLoS Path*., 11(5), e1004806.

Sperschneider J., Gardiner D.M., Dodds P.N., Tini F., Covarelli L., Singh K.B., Manners J.M., Taylor, J.M., 2016. EffectorP: predicting fungal effector proteins from secretomes using machine learning. *New Phyto*., 210(2), 743–761.

Sperschneider J., Williams A.H., Hane J.K., Singh K.B., Taylor, J.M., 2015b. Evaluation of Secretion Prediction Highlights Differing Approaches Needed for Oomycete and Fungal Effectors. *Front. Plant Sci*, 6, 1168.

Stergiopoulos I., Kourmpetis Y.A.I., Slot J.C., Bakker F.T., de Wit P., Rokas A., 2012. In silico characterization and molecular evolutionary analysis of a novel superfamily of fungal effector proteins. *MBE*, 29**(**11), 3371–3384.

Syme R.A., Hane J.K., Friesen T.L., Oliver R.P., 2013. Resequencing and comparative genomics of Stagonospora nodorum: sectional gene absence and effector discovery. *G3: Genes, Genomes, Genetics,* 3(6), 959–969.

Terefe-Ayana D., Yasmin A., Le T.L., Kaufmann H., Biber A., Kuhr A., Linde M., Debener, T., 2011. Mining disease-resistance genes in roses: functional and molecular characterization of the rdr1 locus. *Front. Plant Science*, 2, 35.

Tjalsma H., Bolhuis A., Jongbloed J.D., Bron S., van Dijl J.M., 2000. Signal peptide-dependent protein transport in Bacillus subtilis: a genome-based survey of the secretome. *MMBR*, 64(3), 515–547.

van den Burg H.A, Harrison S.J., Joosten M.H.A.J., Vervoort J., de Wit P., 2006. Cladosporium fulvum Avr4 protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection. *MPMI*, 19(12), 1420–1430.

von Malek B., Debener, T., 1998. Genetic analysis of resistance to blackspot (Diplocarpon rosae) in tetraploid roses. *Theo. & App. Genetics*, 96**(**2), 228–231.

Wang Y., Chen J., Li D.W., Zheng L., Huang, J., 2017. CglCUT1 gene required for cutinase activity and pathogenicity of Colletotrichum gloeosporioides causing anthracnose of Camellia oleifera. *Euro. J. Plant Pathol*., 147(1), 103–114.

Werlemark G., Carlson-Nilsson B.U., Davidson C.G., 2006. Genetic variation in the rose pathogen Marssonina rosae estimated by RAPD. *Int. J. Hortic Sci.*, 12, 63–67.

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., Birch P.R.J., 2007. A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature*, 450**(**7166), 115–118.

Whitaker V.M., Bradeen J.M., Debener T., Biber A., Hokanson S.C., 2010a. Rdr3, a novel locus conferring black spot disease resistance in tetraploid rose: genetic analysis, LRR profiling, and SCAR marker development, Theoretical and applied genetics. *TAG*, 120**(**3), 573–585.

Whitaker V.M., Debener T., Roberts A.V., Hokanson S.C., 2010b. A standard set of host differentials and unified nomenclature for an international collection of Diplocarpon rosae races. *Plant Pathol*., 59**(**4), 745–752.

Whitaker V.M., Hokanson S.C., Bradeen J., 2007. Distribution of rose black spot (Diplocarpon rosae) genetic diversity in eastern north America using amplified fragment length polymorphism and implications for resistance screening. *JASHS*, 132**(**4), 534–540.

Whitaker V.M., Zuzek K., Hokanson S.C., 2007. Resistance of 12 rose genotypes to 14 isolates of Diplocarpon rosae Wolf (rose blackspot) collected from eastern North America. *Plant Breeding*, 126(1), 83–88.

Winnenburg R., Baldwin T.K., Urban M., Rawlings C., Kohler J., Hammond-Kosack K.E., 2006. PHI-base: a new database for pathogen host interactions. *NAR*, 34, D459-D464.

Winnenburg R., Urban M., Beacham A., Baldwin T.K., Holland S., Lindeberg M., Hansen H., Rawlings C., Hammond-Kosack K.E., Kohler J., 2008. PHI-base update: additions to the pathogen host interaction database. *NAR*, 36, D572-D576.

Xiao X., Xie J., Cheng J., Li G., Yi X., Jiang D., Fu Y., 2014. Novel Secretory Protein Ss-Caf1 of the Plant-Pathogenic Fungus Sclerotinia sclerotiorum Is Required for Host Penetration and Normal Sclerotial Development. *MPMI*, 27**(**1), 40–55.

Yang F., Jensen J.D., Svensson B., Jorgensen H.J.L., Collinge D.B., Finnie, C., 2012. Secretomics identifies Fusarium graminearum proteins involved in the interaction with barley and wheat. *Mol. Plant. Pathol*., 13(5), 445–453.

Yin Y., Mao X., Yang J., Chen X., Mao F., Xu Y., 2012. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *NAR*, 40, W445-451.

Zhu S., Cao Y.Z., Jiang C., Tan B.Y., Wang Z., Feng S., Zhang L., Su X.H., Brejova B., Vinar T., Xu M., Wang M.X., Zhang S.G., Huang M.R., Wu R., Zhou, Y., 2012 Sequencing the genome of Marssonina brunnea reveals fungus-poplar co-evolution, *BMC genomics*, 13, 382.

5. Interaction of roses with a biotrophic and a hemibiotrophic leaf pathogen leads to differences in defense transcriptome activation

Enzo Neu[1.], Helena Sophia Domes[1.], Helgard Kaufmann[1.], Marcus Linde[1.], Ina Menz[1.], Thomas Debener[1.*]

[1.] Department of Molecular Plant Breeding, Institute for Plant Genetics, Leibniz Universität Hannover, Hannover 30419, Germany

Type of authorship:	First author
Type of article:	Research article
Contribution to the article:	Planned and performed inoculation experiments with *D. rosae* and *P. pannosa.* Involved in performing the experiments to validate the transcriptomic data via qPCR Analysis of the MACE-data. Wrote majority parts of the manuscript. Prepared the majority of tables and figures.
Contribution of the other authors:	Helgard Kaufmann generating the samples for MACE-analysis, which were inoculated with *P. pannosa.* Helena Domes and Ina Menz contributed substantial in performing the experiments to validate the transcriptomic data via qPCR and wrote parts of the manuscript. Helena Domes and Ina Menz contributed in analysing and interpretation of the transcriptomic data. All authors contributed in revision of the manuscript.

Journal:	Plant Molecular Biology
Date of submission:	09.04.2018
Impact factor:	3.356

Interaction of roses with a biotrophic and a hemibiotrophic leaf pathogen leads to differences in defense transcriptome activation

Enzo Neu[1.], Helena Sophia Domes[1.], Helgard Kaufmann[1.], Marcus Linde[1.], Ina Menz[1.], Thomas Debener[1.*]

[1.] Department of Molecular Plant Breeding, Institute for Plant Genetics, Leibniz Universität Hannover, Hannover 30419, Germany

* **Correspondence:**

Thomas Debener

debener@genetik.uni-hannover.de

Key Message: Transcriptomic analysis resulted in up-regulation of genes related to common defense mechanisms for black spot and in down-regulation of genes related to photosynthesis and cell wall modification for powdery mildew.

Keywords: black spot, powdery mildew, MACE analysis, high-throughput qPCR, WRKY genes, PR genes

Manuscript length: 7575 words

Abstract length: 224

Tables: 3

Figures: 10

Abstract

Biotrophic pathogens establish a long-term feeding relationship with their hosts, while the lifestyle of hemibiotrophic pathogens changes from a biotrophic stage to a necrotrophic stage, in which the host cells are killed.

We analyzed the leaf transcriptome of roses after inoculation with either the obligate biotrophic fungus *Podosphaera pannosa*, which causes powdery mildew, or the hemibiotrophic fungus *Diplocarpon rosae*, which causes black spot. Transcriptomic changes were analyzed using next-generation sequencing based on the MACE (massive analysis of cDNA ends) approach, which was validated by high-throughput qPCR. Using highly stringent criteria, we identified a large number of differentially regulated genes. Surprisingly, most of the differentially regulated genes were specific to interactions with either *P. pannosa* or *D. rosae*. A common set of the differentially regulated genes comprised pathogenesis-related (PR) genes, such as homologues of PR10, chitinases and defense-related transcription factors, such as various WRKY genes, indicating a conserved PTI (PAMP- (pathogen associated molecular pattern) triggered immunity) reaction. Specific regulation in response to the hemibiotrophic black spot fungus was detected not only for genes from the phenylpropanoid and flavonoid pathways but also for individual PR genes, such as paralogues of PR1 and PR5 and other factors of the salicylic acid signaling pathway. In contrast, inoculation with *P. pannosa* did not result in the specific up-regulation of functional groups related to common defense mechanisms; rather, it resulted in the down-regulation of genes related to photosynthesis and cell wall modification.

Introduction

Fungal pathogens have developed different lifestyles to colonize their hosts. Biotrophic fungi develop nutrient-absorbing structures, such as haustoria, to establish a long-term feeding relationship without killing their hosts. In contrast, necrotrophic fungi kill the host cells, often by secreting toxins, to feed on the dead tissue, while hemibiotrophic pathogens often change their lifestyle during development from an early biotrophic to a later necrotrophic stage (Horbach et al., 2011).

In their natural environment, roses are confronted with many different pathogens. Among them, the biotrophic fungus *Podosphaera pannosa*, which causes powdery mildew, and the hemibiotrophic *Diplocarpon rosae*, which causes black spot, are the most common and damaging pathogens in roses cultivated in the greenhouse and the field, respectively (Debener and Byrne, 2014).

The life cycle of *D. rosae* begins with the germination of its conidia within eight to nine hours after inoculation, and three hours later, appressoria may already be formed. After an additional three hours, subcuticular hyphae and first haustoria may be developed (Aronescu, 1934). In this biotrophic phase, all fungal structures in the host cells are surrounded by the plasma membrane of the host. After approximately 6 dpi (days post inoculation), necrotrophic intracellular hyphae, growing from cell to cell without a surrounding host membrane, may be observed, leading to the formation of acervuli in the necrotrophic stage. The resulting circular black spots that are typical infection symptoms are mostly surrounded by living host cells, so-called “green islands” (Gachomo et al., 2006). Few details about the infection process of *P. pannosa* in roses are available. Germ tubes may be observed 2 to 6 hpi (hours post inoculation), rapidly developing into a mycelium growing on the leaf surface. After approximately 1 dpi, multilobed haustoria develop under the appressoria, which are formed in regular distances (Linde and Shishkoff, 2003). Under favorable conditions, asexually produced spores are released after 3 to 5 days, completing the life cycle on the living host tissue (Coyier, 1983).

Research on these two diseases mainly focuses on the identification and characterization of resistance gene loci. Thus, a number of *R*-genes (resistance genes) against *D. rosae* have been characterized (von Malek and Debener, 1998, Hattendorf et al., 2004, Whitaker et al., 2010a), including one TNL-type gene, *Rdr1*, that mediates a broad-spectrum resistance against different black spot isolates (Kaufmann et al., 2003, Terefe-Ayana et al., 2011, Menz et al., 2017). On the pathogen side, at least 11 pathogenic races have been described by means of different sets of host plants, and a draft genome of the isolate DortE4 is available (Neu et al., 2017, Whitaker et al., 2010b).

For the resistance against powdery mildew, three major resistance genes have been identified, and several QTL regions have been mapped (Linde et al., 2004; Xu et al., 2005, Dugo et al., 2005, Linde et al., 2006, Hosseini Moghaddam et al., 2012). In

addition, four homologs to the mildew resistance locus o (MLO), which mediates broad-spectrum resistance via loss-of-function mutations in other species, have been mapped and characterized (Kaufmann et al. 2012). For *P. pannosa*, up to eight pathogenic races and many different isolates have been described, but no sequence information is available, except for some ITS sequences (Linde and Debener, 2003, Leus et al., 2006).

The recognition of a pathogen can either be mediated by pattern-recognition-receptors in the cell membrane, which recognize highly conserved pathogen- or microbe-associated molecular patterns (PAMP or MAMP) on the surface of a pathogen, e.g., fungal chitin or bacterial flagellin, or by R proteins recognizing the effectors secreted from the pathogen into the host cell. These two mechanisms are called PAMP-triggered immunity (PTI) and effector-triggered immunity (ETI), and the latter is often accompanied by a so-called “hypersensitive response” (Bent and Mackey, 2007, Jones and Dangl, 2006).

The recognition of a pathogen by plant cells leads to drastic changes in the transcriptome and activates a diverse set of immune responses, including the synthesis of secondary metabolites, cell wall modifications and the expression of pathogenesis-related (PR) genes (Slusarenko et al., 2000). Plant hormones play an important role as signaling molecules in plant defenses. Salicylic acid (SA) is involved in the defense response to biotrophic and hemibiotrophic pathogens, whereas jasmonate (JA) and ethylene (ET) are the main hormones involved in the response to necrotrophic pathogens. Both pathways have often been described as antagonistic. In addition, cross-talk with other phytohormones can modulate the responses to pathogens/stress (Derksen et al., 2013).

In this study, we compared changes in the leaf transcriptome of roses induced by the hemibiotrophic fungus *D. rosae* and the biotrophic fungus *P. pannosa* by using the MACE technique to analyze the defense mechanisms of roses against both pathogens. Data from three independent inoculation experiments comprising samples from three time points in the early stages of infection were stringently analyzed to provide comprehensive insights into the two pathosystems.

Material and Methods

Plant Material

For the MACE analysis and the multiplication of the fungal isolates, the susceptible rose cultivar ‘Pariser Charme’ (PC) was propagated in vitro, as previously described (Davies, 1980, Debener T. et al., 1998, Linde and Debener, 2003, Dohm et al., 2001), rooted, transplanted into fertilized Einheitserde T substrate (Einheitserdewerke Gebr. Patzer, Sinntal-Altengronau, Germany) in 9-cm pots and cultivated under semi-controlled conditions (12 h light/12 h darkness). The plants used for the validation of the MACE analysis were cultivated in climate chambers under short-day conditions (8 h light/16 h darkness) at 22 °C.

Disease assays

For the black spot inoculations, the *D. rosae* isolate DortE4 was used. The single conidial isolate was maintained on detached leaves of the susceptible cultivar PC, as previously described (von Malek and Debener, 1998). Young leaves were infected with a suspension of 500000 conidia/mL with a vaporizer and kept on moist tissue paper in translucent plastic boxes in an air-conditioned laboratory at 20 °C.

A multispore isolate of powdery mildew originating from the host genotype PC, which was grown in the greenhouses at Leibniz Universität Hannover (Germany), was used for inoculation. Very young, recently unfolded PC leaves were placed into an infection box covered with a 14 cm height 100 µm nylon mesh. Infected leaves were rubbed over the mesh to cover the leaves with conidia. The conidia were allowed to settle for 30 min to ensure that all conidia reached the leaf surfaces. To estimate the density of the conidia, a microscope slide was placed next to the leaves, and conidia were counted under a microscope at 100-fold magnification. A minimum density of 16 conidia/mm^2 was used. Following this, infected leaves were stored in translucent plastic boxes on moist tissue paper at 20 °C. Uninoculated leaves were used as controls and stored in identical conditions as the inoculated samples. The treatments and infection time points for both pathogens are listed in Table 1. For each time point, three completely independent inoculation experiments were conducted for the MACE experiment, and independent sets of three additional experiments were conducted

for the qPCR experiment. These experiments were treated as biological repeat experiments.

Table 1: Overview of the treatments and infection time points.

Treatment	0 hpi[1]	24 hpi	72 hpi
Powdery mildew	PC[2]+PP[3] 0 hpi	PC+PP 24 hpi	PC+PP 72 hpi
Black spot	PC+DR[4] 0 hpi	PC+DR 24 hpi	PC+DR 72 hpi
Control (no inoculation)	PC-Co[5] 0 hpi	PC-Co 24 hpi	PC-Co 72 hpi

[1]Hours post inoculation, [2]rose genotype 'Pariser Charme', [3] *Podosphaera pannosa*, [4] *Diplocarpon rosae*, [5] control

Microscopic analysis

For the microscopic analysis, the control leaves and leaves infected with either *D. rosae* or *P. pannosa* were sampled at 0, 24, and 72 hpi. Leaf pieces of approximately 1 cm^2 were fixed, stained with Alexa Fluor 488 conjugated wheat germ agglutinin (Invitrogen, Carlsbad, USA) and examined as previously described (Menz et al., 2017).

MACE

Transcriptomic data were generated by using the MACE technique (Kahl et al., 2012). A particular feature of this technique is that only one sequence (tag) per cDNA molecule is generated, so normalization to the length of the respective transcript/gene model is not necessary. RNA for this analysis was extracted from the three independent biological replicates with the RNeasy® Plant Mini Kit from Qiagen (Hilden, Germany) following the manufacturer's instructions, with an additional DNase digestion step via the Ambion VR DNA-free™ Kit from Life Technologies (Carlsbad, CA, USA) to remove the remaining genomic DNA. The cDNA synthesis

and sequencing was performed by GenXPro (Frankfurt am Main, Germany). The sequencing data were already quality- and adapter-trimmed by the provider. The raw reads of this study have been placed in the NCBI Sequence Read Archive (SRA) under the accession numbers SRR6879138 to SSR6879164. Further processing of the sequences was performed by using CLC Genomic Workbench 9.0.1 (Qiagen, Hilden, Germany). The remaining pieces of the poly-A tail were removed by an additional adapter trimming step using a polyA-adapter. Reads shorter than 35 bp were discarded. The trimmed sequences were mapped to the genomic sequence of *Rosa chinensis* var. 'Old Blush' (Hibrand et al., 2018) using the following parameters: mismatch cost=2, insertion cost=3, deletion cost=3, length fraction=0.9, similarity fraction=0.9, strand specific=both, maximum number of hits=10 and expression value=unique counts. The expression values were normalized by means of tags per million (TPM) normalization and were log2 transformed. Differential gene expression was analyzed with the EdgeR package (Robinson et al., 2010) with a false discovery rate (FDR) adjustment.

Only genes detected in all three repeat experiments of at least one condition that had at least 1 TPM (tag per million reads) and a minimum mean fold change +/- 3-fold, with an FDR-adjusted p value ≤ 0.05, were considered to be differentially expressed.

Additional data analysis

In addition to the annotation presented by Hibrand et al. (2018), Blast2GO 4.1.9 software (Conesa et al., 2005) was used to classify sequences with GO terms and generate an automatic functional description derived from a BlastP search (E-Value cut-off 1E-10) against a subset of the NR database that only contained plant sequences (Ashburner et al., 2000). These GO terms were used for an enrichment analysis of the differentially expressed genes (DEGs) by means of Fisher's exact test implemented in Blast2GO (Conesa et al., 2005).

For the visualization of the regulated genes and pathways, we used MapMan 3.6.0 software (Thimm et al., 2004) with a customized mapping file of all DEGs generated with the Mercartor webserver (Lohse et al., 2014) using default settings.

Based on the normalized expression values, the Spearman rank correlations between samples and biological repeat experiments were calculated in R 3.4.0 (R Core Team).

Validation of MACE analysis using high-throughput qRT-PCR

For the validation of the MACE analysis, an independent set of three inoculations representing independent biological replicates was performed. From 20–30 mg of infected leaf tissue (0 hpi, 24 hpi and 72 hpi), the total RNA was isolated using the Quick-RNA™ MiniPrep Plus kit from Zymo Research (Irvine, USA) in accordance with the manufacturer's protocol, with some minor modifications. The leaf material was frozen in liquid nitrogen and disrupted using a bead mill. Dithiothreitol (DTT) was added to the lysis buffer to a final concentration of 50 mM.

For the cDNA synthesis, 500 ng of total RNA was processed with the High Capacity cDNA Reverse Transcription Kit from Applied Biosystems VR (Carlsbad, USA) according to the manufacturer's instructions. For the high-throughput qPCR (BioMark), a set of 28 genes was chosen among the differentially regulated genes identified by the MACE analysis. Primers for the genes were constructed with Primer3plus (Rozen and Skaletsky, 2000) and are listed in Supplementary Table 1. The primer efficiency was tested with a dilution series (1:4, 1:16, 1:64, 1:256) using the StepOnePlus™ system from Applied Biosystems (Austin, USA), as described by Menz et al. (2017). The expression of the differentially regulated genes and the three reference genes (*TIP, SAND* and *UBC*) (Klie and Debener, 2011) were analyzed on a Fluidigm Dynamic Array IFC (96.96) following the manufacturer's instructions. The specific target amplifications (STA) were diluted fivefold, and the qPCR conditions were as follows: 60 s at 95 °C, 30 cycles of 95 °C for 5 s and 20 s at 64 °C, and a final melting curve analysis. Data were processed with Fluidigm Real-Time PCR Analysis Software (4.3.1). A quality threshold of 0.65, a linear baseline correction and an auto global cycle threshold (Ct) were used. Ct values were used to calculate expression ratios with the REST 2009 software (V2.0.13, Qiagen, Hilden, Germany). The expression ratios were log2 transformed, and the coefficient of correlation (Pearson) of the significantly up- or down-regulated genes in both the MACE and BioMark analysis were calculated in R 3.4.0 (R Core Team).

Results

To characterize the responses of the rose leaf transcriptome during compatible interactions of roses with the hemibiotrophic fungus *D. rosae* and the obligate biotrophic fungus *P. pannosa*, inoculation experiments were performed for three different time points (Table 1).

Successful infection was confirmed by the appearance of disease symptoms, and the microscopic evaluation of the early time points was performed in parallel with the tissue sampling for RNA extraction. The development of *D. rosae* and the development of *P. pannosa* were similar at the beginning. At 24 hpi, the spores germinated and the first haustoria formed (Fig. 1. B and F). At 72 hpi, *D. rosae* developed long-range hyphae, and numerous haustoria were formed in the epidermis and the underlying mesophyll layer (Fig. 1. C-D). However, the development after 72 hpi was more progressed in *P. pannosa* than in *D. rosae.* More fungal hyphae spread across the wide parts of the leaf, and a high number of haustoria formed in the epidermis of the plant (Fig. 1. G-H).

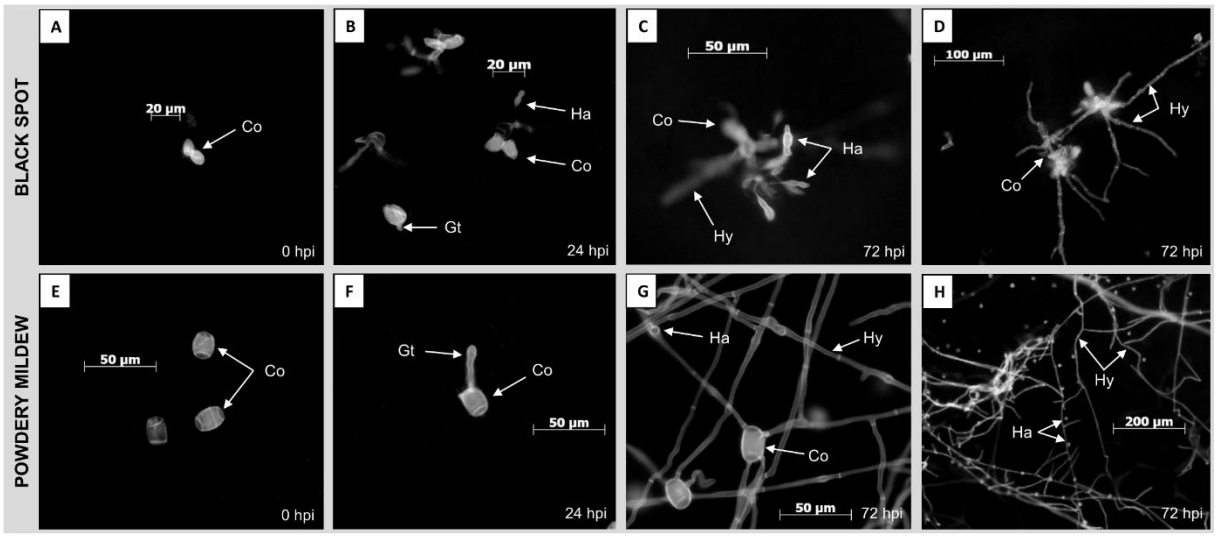


Fig.1: Microscopic analysis of the two interaction systems. Interaction between the susceptible genotype PC and *D. rosae* (A-D) and *P. pannosa* (E-H) at different time points (0 hpi (A, E), 24 hpi (B, F) and 72 hpi (C, D, G, H). Co: conidia; Gt: germ tube; Ha: haustoria; Hy: hyphae. Samples were stained with Alexa Fluor 488 conjugated wheat germ agglutinin (WGA).

Analysis of gene expression by MACE sequencing

MACE sequencing was performed for the inoculated and control samples for all time points (0, 24, 72 hpi) in the three biological repeats, each represented by an independent inoculation experiment. The sequencing resulted in 6 to 30 million high-quality reads per library, with an average output of 12.7 million reads per library (Table 2). Between 79.2% and 94.2% of these reads could be mapped to the genome sequence of the 'Old Blush' rose variety (Hibrand et al., 2018). The majority of the sequenced reads (76.3-92.9%) mapped uniquely to the genome.

To get an overview of the expression level relationships between the samples, the Spearman correlation was calculated (Fig. 2). Three larger clusters were visible. One contained all samples at 0 hpi, except for the first biological repeat of the inoculation with *D. rosae* (BR1_PC+DR_0). This sample fell into the second cluster with the remaining control samples (PC-Co_24 and PC-Co_72) and all the biological repetitions of the inoculation with *P. pannosa* at 24 hpi (PC+PP_24). The third cluster comprised samples from the inoculation with *P. pannosa* at 72 hpi (PC+PP_72) and samples from the inoculation with *D. rosae* at 24 and 72 hpi (PC+DR_24 and PC+DR_72). Noticeably, in each cluster, the samples of the third biological repeat formed subclusters.

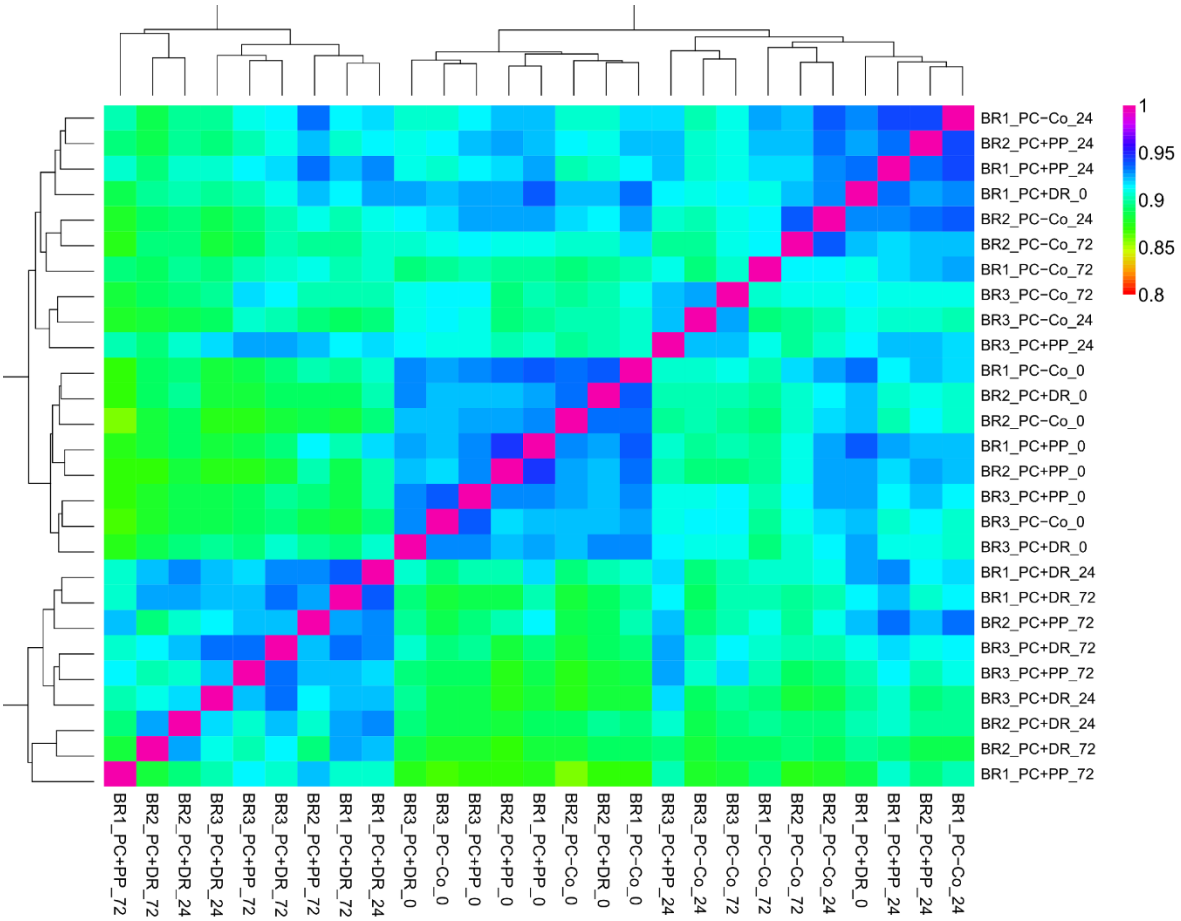


Fig. 2: Heat map of the Spearman rank correlations between samples and biological repeat experiments for the normalized expression values. The nomenclature of the experiments (Co = control, PP= *P. pannosa*, DR = *D. rosae*, BR1-3 = biological repeats) is the same as that in Table 2.

Table 2: Summary of MACE sequencing and read mapping. The treatments of the genotype PC with *D. rosae* (DR), *P. pannosa* (PP) and the control (Co) at different time points are listed. BR1-BR3: biologically repeated experiments.

Sample	Biological Repetition	No. of Reads	Average length [bp]	Total reads mapped [No. | %]	Uniquely mapped reads [No. | %]
PC-Co 0 hpi	BR 1	15,520,605	89.6	14,601,985 | 94.08	14,341,828 | 92.41
	BR 2	13,220,333	89.6	12,457,221 | 94.23	12,264,139 | 92.77
	BR 3	6,930,127	109.1	6,387,578 | 92.17	6,270,867 | 90.49
PC-Co 24 hpi	BR 1	14,359,485	89.1	13,251,849 | 92.29	12,911,387 | 89.92
	BR 2	14,581,728	89.7	13,759,727 |94.36	13,489,985 | 92.51
	BR 3	7,036,752	109.5	6,452,053 | 91.69	6,324,465 | 89.88
PC-Co 72 hpi	BR 1	5,963,608	89.1	5,561,118 | 93.25	5,424,380 | 90.96
	BR 2	10,517,799	89.5	9,951,615 | 94.62	9,765,980 | 92.85
	BR 3	8,811,940	110.3	8,074,847 | 91.64	7,916,510 | 89.84
PC+DR 0 hpi	BR 1	20,487,155	89.0	19,197,580 | 93.71	18,828,415 | 91.90
	BR 2	12,775,434	89.6	12,007,935 | 93.99	11,807,189 | 92.42
	BR 3	12,254,300	110.7	11,234,434 | 91.68	11,050,633 | 90.18
PC+DR 24 hpi	BR 1	18,202,210	88.8	16,860,165 | 92.63	16,470,501 | 90.49
	BR 2	8,899,047	89.2	8,276,781 | 93.01	8,083,119 | 90.83
	BR 3	7,898,427	111.0	7,214,627 | 91.34	7,036,258 | 89.08
PC+DR 72 hpi	BR 1	19,632,628	89.3	17,721,058 | 90.26	17,284,909 | 88.04
	BR 2	8,916,530	89.4	7,787,907 | 87.34	7,606,336 | 85.31
	BR 3	9,121,966	110.5	8,186,839 | 89.75	7,991,056 | 87.60
PC+PP 0 hpi	BR 1	29,689,430	88.6	27,296,193 | 91.94	26,730,856 | 90.03
	BR 2	16,708,034	89.8	15,517,322 | 92.87	15,221,035 | 91.10
	BR 3	9,343,766	109.6	8,569,849 | 91.72	8,405,753 | 89.96
PC+PP 24 hpi	BR 1	16,190,796	88.7	14,906,067 | 92.07	14,525,503 | 89.71
	BR 2	11,077,018	89.2	10,320,782 | 93.17	10,080,751 | 91.01
	BR 3	6,919,764	110.0	6,243,326 | 90.22	6,089,403 | 88.00
PC+PP 72 hpi	BR 1	7,128,516	89.1	5,644,136 | 79.18	5,439,140 | 76.30
	BR 2	22,205,641	88.6	19,480,993 | 87.73	18,876,738 | 85.01
	BR 3	8,510,484	109.9	7,353,280 | 86.40	7,160,598 | 84.14

Validation of MACE analysis by high-throughput RT-qPCR

To validate the results of the MACE analysis, three additional inoculation experiments were analyzed by a high-throughput RT-qPCR system. For this purpose, a set of 28 significantly up- and down-regulated genes was chosen based on the MACE results. The selection mainly focused on genes that showed similar differences in expression between the infected and non-infected samples in all three biological replicates. A close correlation (R=0.82) between the significant log2-fold changes in the MACE and RT-qPCR results was detected, and except for three data points that were down-regulated in the qPCR results and up-regulated in the MACE results (left quadrant), the same expression trends were observed (Fig.). The selected genes, primer sequences, PCR amplification efficiencies and expression data of the MACE and RT-qPCR analysis are listed in Supplementary Table 1.

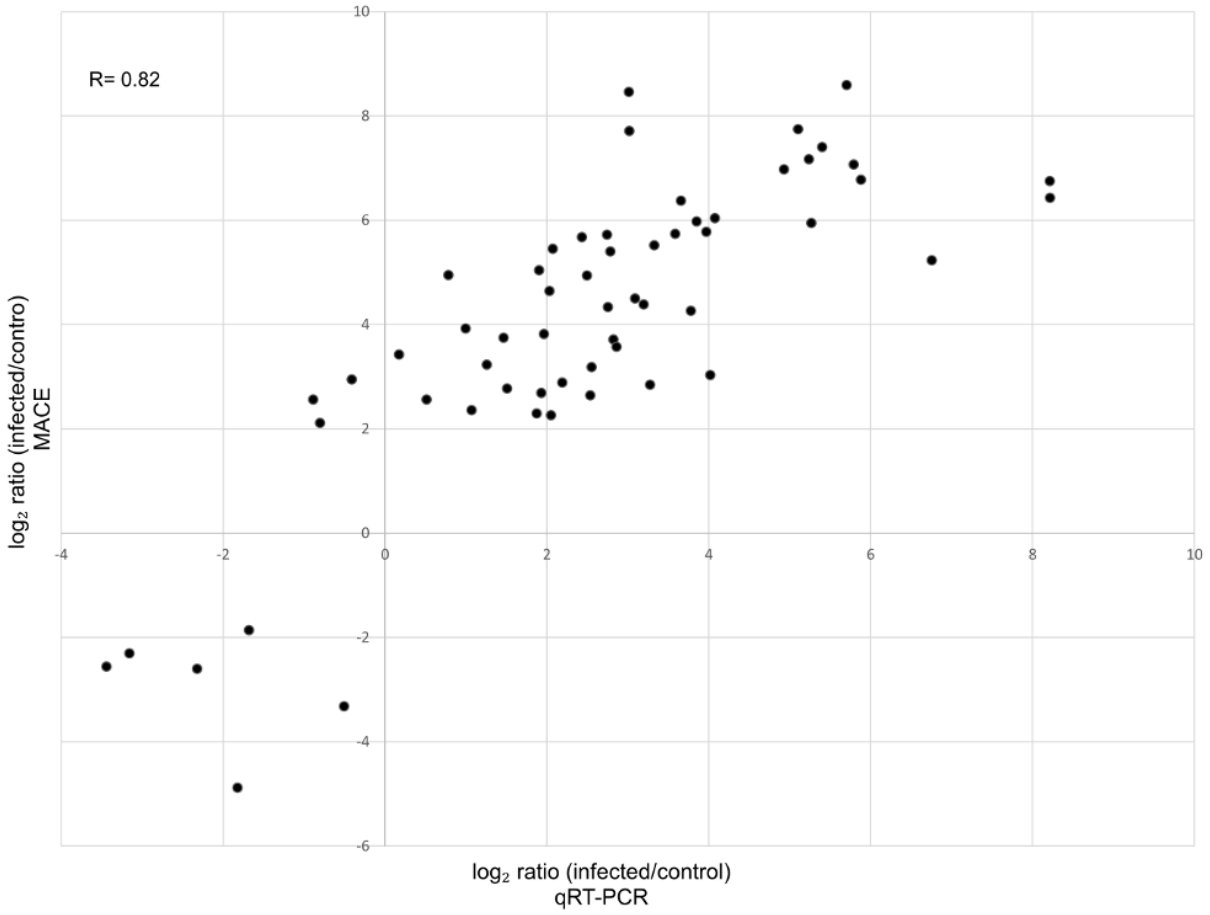


Fig. 3: Scatter plot for 28 genes analyzed with both MACE and RT-qPCR at three time points. Only the expression ratios of the significantly up- or down-regulated genes from the MACE and RT-qPCR data are depicted, resulting in 58 data points. Pearson's correlation coefficient is also shown.

Response of roses to powdery mildew infection

The response of PC to infection with *P. pannosa* leads to major changes in the leaf transcriptome. Supplementary Table 2 contains normalized tag counts as expression values, the results of the EdgeR statistical analysis and the corresponding annotations of the comparison between the *P. pannosa*-inoculated samples with the control for all three time points. In total, 1450 genes (650 up- and 800 down-regulated) were differentially expressed under the stringent criteria applied in this analysis (Fig. 4). The majority of these genes, representing approximately 80% of the up- and down-regulated genes, were differentially expressed at 72 hpi. The smallest number of differentially expressed genes were detected at 24 hpi. The majority of the up- or down-regulated genes at 24 hpi are also regulated at 72 hpi. Directly after the inoculation (0 hpi), 131 genes were up-regulated and 142 genes were down-regulated in the inoculated leaves compared to the control samples, but only a few of the changes were also observed at the other time points.

To get an overview of the biological responses linked to the differentially expressed genes, we performed a GO enrichment (Supplementary Table 3) analysis using Blast2GO (Conesa et al., 2005) and the MapMan tool (Thimm et al., 2004), with a particular interest in better visualizing the defense-related mechanisms. The GO terms enriched in the set of up-regulated genes (72 hpi) indicate that a defense response to the fungus was in progress, as shown by their enrichment in GO terms such as “defense response”, “response to biotic stimulus” “chitin catabolic process” and “chitinase activity”. In accordance with the drastic changes in the transcriptome, the GO term “transcription factor activity, sequence-specific DNA binding” was also overrepresented in the up-regulated sequences. In addition, GO terms involved in the regulation of peptidases were noticeably enriched. The processes represented by the significantly overrepresented GO terms in the set of down-regulated sequences were diverse, but many of the enriched GO terms were related to photosynthesis (e.g., “photosynthesis”, “photosynthesis, light reaction”, “chlorophyll binding”, “response to light stimulus”) or cell wall organization (e.g., “cell wall organization or biogenesis”, “cell wall biogenesis”, “cellular polysaccharide metabolic process”, “pectin catabolic process”). In addition, the auxin-mediated signaling pathway seems to have been down-regulated (“auxin-activated signaling pathway”, “cellular response to auxin stimulus”).

Fig. 5 displays the responses to biotic stress visualized by the MapMan analysis. As the figure shows, the majority of the response genes were down-regulated. In particular, many genes involved in cell wall modification and secondary metabolism were down regulated. Additionally, a large proportion of genes involved in signaling were down-regulated, whereas others were strongly up-regulated. Some PR proteins, as well as genes involved in all hormone signaling pathways, were regulated. Notably, some WRKY transcription factors, ethylene response factors (ERFs), peroxidases and glutathione-S-transferases were strongly up-regulated.

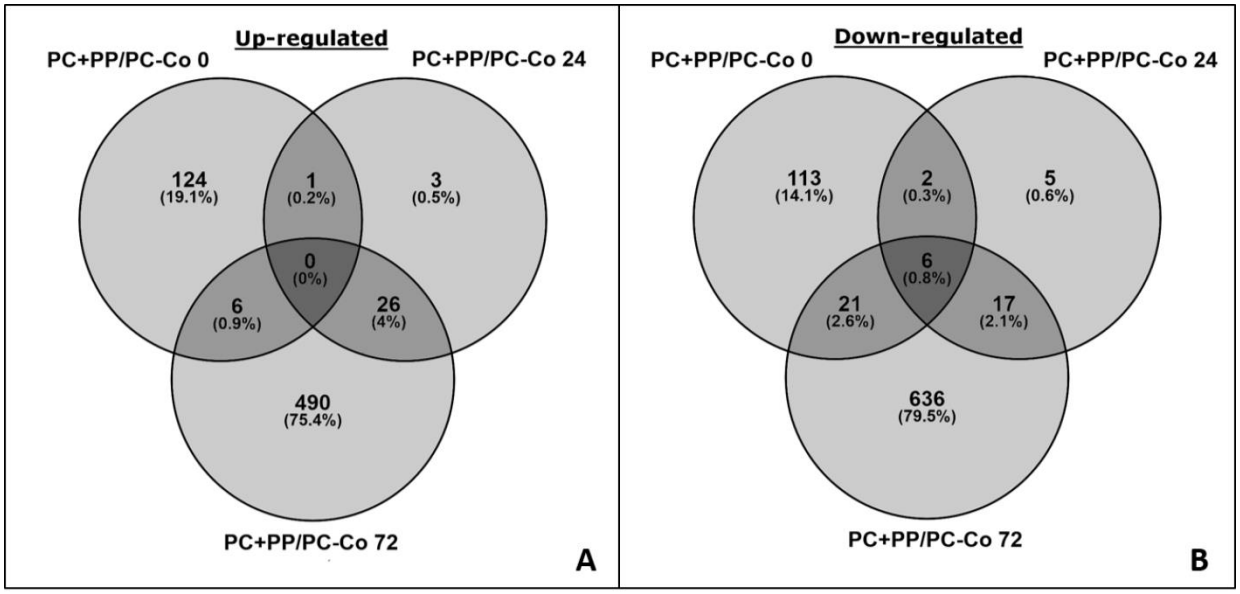


Fig. 4: Venn diagram of significantly regulated genes after *P. pannosa* inoculation. Genes, that were significantly up-(A) and down-regulated (B) compared to the control samples at different time points (0, 24, 72 hpi) are shown.

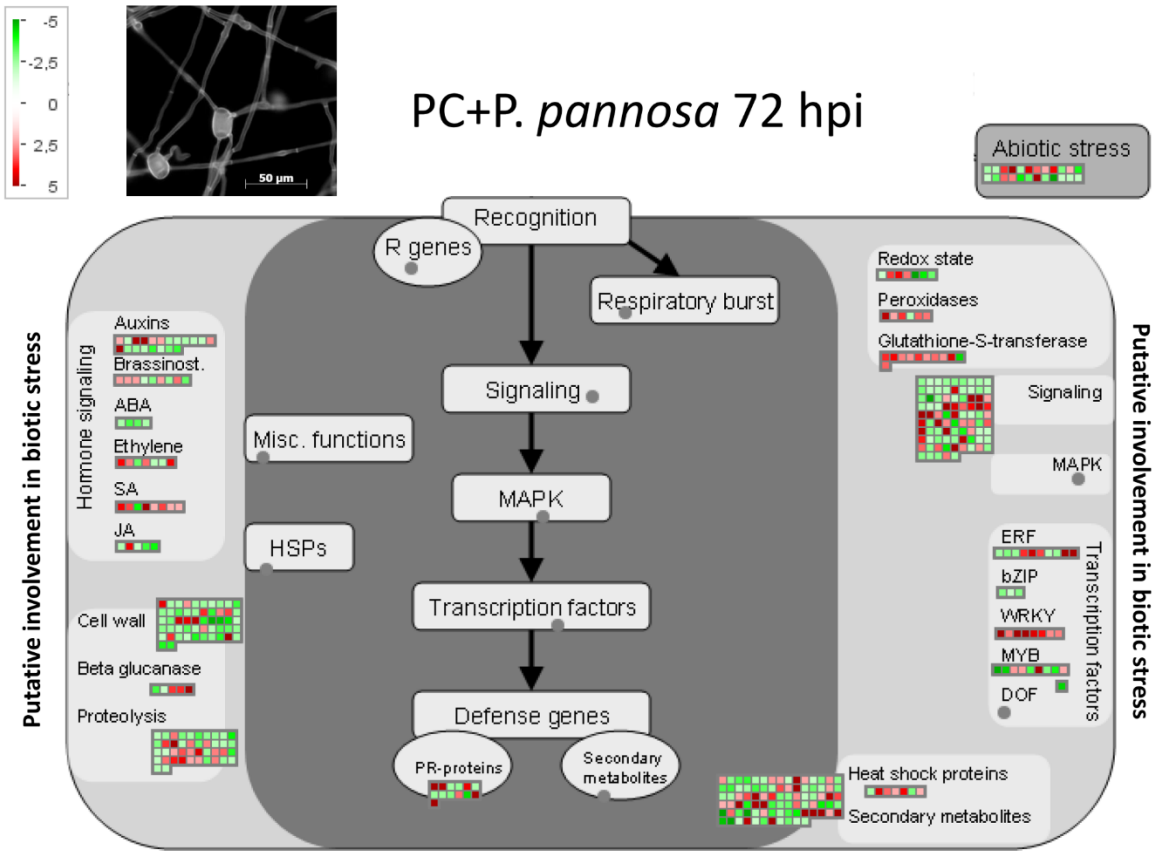


Fig. 5: Overview of the regulated processes involved in the response to biotic stress visualized by MapMan. Genes that were up-regulated in the inoculated samples compared to the control samples are indicated in red, and the down-regulated genes are indicated in green. Only significant changes in expression are displayed. The scale bar represents log2-fold changes (PC+PP/PC-Co. 72 hpi).

Response of roses to black spot infection

Major changes in the leaf transcriptome also occurred in response to *D. rosae* inoculation. The expression data in the form of normalized tag counts, the results of the EdgeR statistical analysis and the corresponding annotations for the comparison of the *D. rosae*-inoculated samples with the control samples for all three time points are presented in Supplementary Table 4.

Almost 2000 genes changed their expression, and the majority (1158) of these were up-regulated, with only 828 that were down-regulated (Fig. 6). Less than 100 genes were considered differentially expressed directly after the inoculation at 0 hpi. Approximately 80% of both the up- or down-regulated genes were observed at 24

hpi. The majority of the up-regulated genes and many of the down-regulated genes show a similar regulation pattern at 72 hpi.

Similarities between the two time points were also visible in the enrichment analysis (Supplementary Table 3), where the majority of the overrepresented GO terms were the same for both 24 and 72 hpi. In the reaction to *D. rosae*, GO terms related to the defense response (e.g., “response to biotic stimulus“, “response to fungus”, “defense response”, “chitinase activity”) were highly enriched in the set of up-regulated genes. In addition, GO terms such as “calcium ion binding”, “kinase activity” or “reactive oxygen species metabolic process”, which are involved in signaling processes, were also significantly overrepresented, as well as GO terms representing secondary metabolism, especially those related to the phenylpropanoid and flavonoid pathways (“secondary metabolite biosynthetic process”, “phenylpropanoid metabolic process”, “caffeate O-methyltransferase activity”, “flavonoid metabolic process”). Interestingly, the GO term “transcription factor activity, sequence-specific DNA binding” was enriched in both the up- and down-regulated gene sets. In addition, GO terms representing photosynthesis-related (“photosynthesis, light harvesting”, “chlorophyll binding”) and cell wall organization-related mechanisms (“pectin metabolic process”, “cellular polysaccharide catabolic process”) were enriched in the down-regulated genes. The GO-enrichment also indicates a down-regulation of the auxin-mediated signaling pathway (“auxin-activated signaling pathway”, “cellular response to auxin stimulus”).

The MapMan analysis (Fig. 7) shows the up-regulation of many genes in response to *D. rosae*, especially genes such as kinases involved in signaling processes. In addition, many PR genes, WRKY transcription factors, peroxidases and glutathione-S-transferases seemed to be strongly up-regulated. Compared to the reaction to *P. pannosa*, there were more regulated genes related to the hormone signaling pathway, particularly the ethylene pathway, and secondary metabolism. In addition, there were more and stronger regulations of PR-proteins. Many genes involved in cell wall modification were mainly down-regulated, as well as heat shock proteins and genes related to abiotic stress.

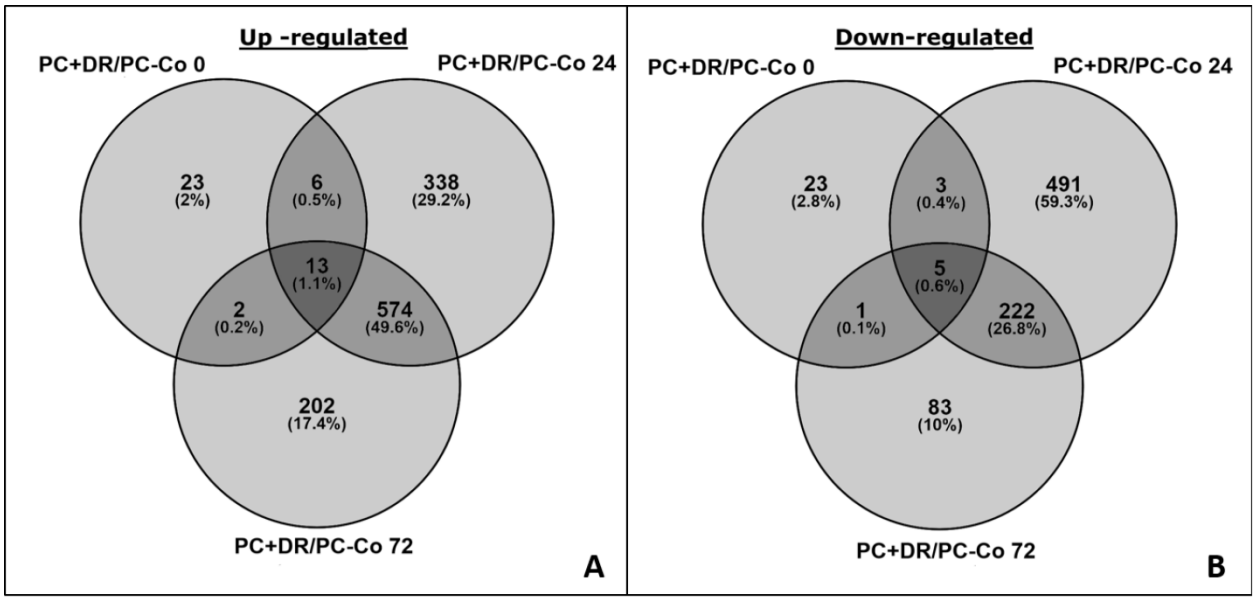


Fig. 6: Venn diagram of significantly regulated genes after *D. rosae* inoculation. Genes, that were significantly up-(A) and down-regulated (B) compared to the control samples at different time points (0, 24, 72 hpi) are shown.

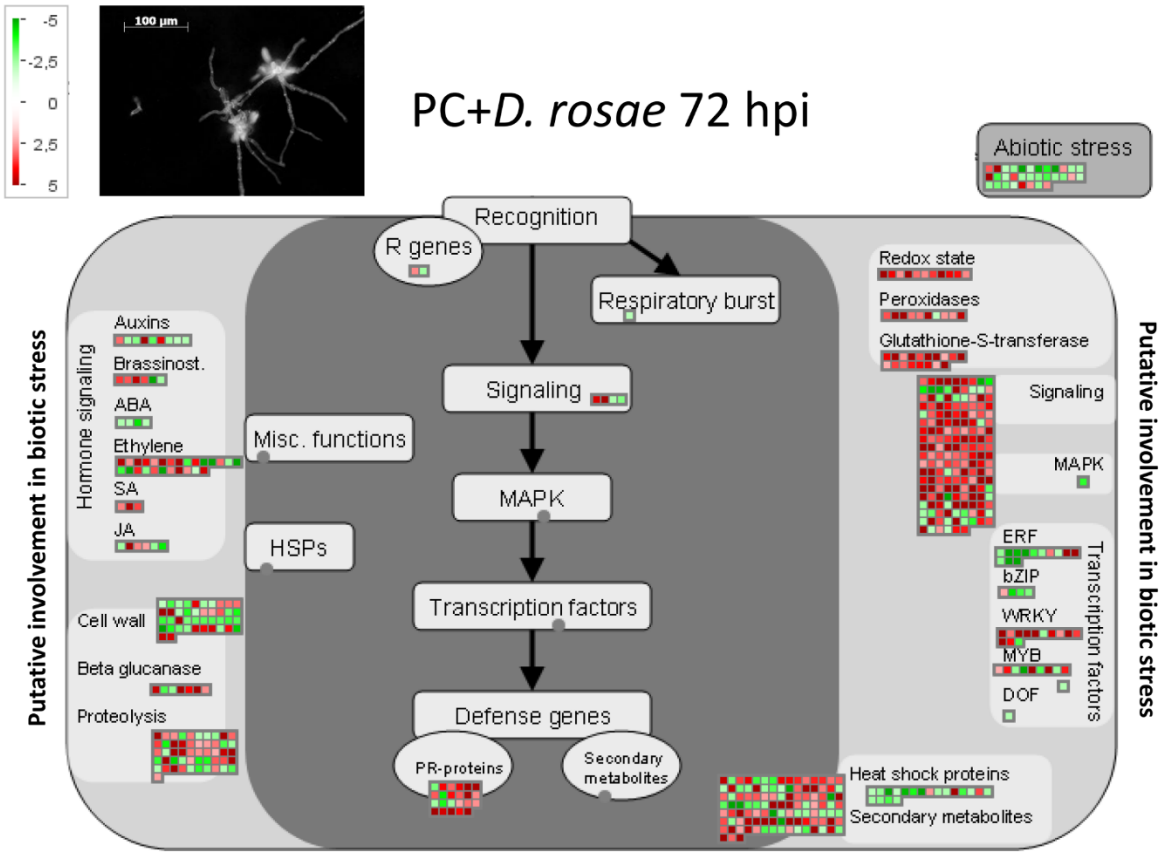


Fig. 7: Overview of the regulated processes involved in the response to biotic stress visualized by MapMan. The genes up-regulated in the inoculated samples compared to the control samples are indicated in red, and the down-regulated genes are indicated in green. Only significant changes in expression are displayed. The scale bar represents log2-fold changes (PC+DR/PC-Co. 72 hpi).

Comparison of the response of roses to both pathogens

Based on the comparison of the responses of roses to both pathogens, the majority of the significantly regulated genes were specific to the reaction to one of the pathogens. Only a smaller number of the genes were regulated in response to both pathogens. One of the major differences between the two responses is the time point when the genes were regulated. During the interaction with *D. rosae,* many of the responses occurred at 24 hpi (Table 2, Fig. 6). In contrast, almost no change in the gene expression was visible in the interaction with *P. pannosa* at 24 hpi (Fig. 4). In the comparison of the later responses at 72 hpi (Fig. 8), it is noticeable that only approximately 25% of the significantly up-regulated genes and 21.6% of the significantly down-regulated genes were identical between the different pathogen interaction systems. The remaining genes were exclusively regulated in response to only one pathogen. The most strongly differentially regulated genes are listed in Table 3. These genes comprised some defense-related genes, such as the major allergen

Pru ar 1, belonging to the PR10 family. Furthermore, some genes involved in signaling pathways, such as kinases, transcription factors and a calmodulin, as well as genes encoding cytochrome P450 proteins, were among the most highly regulated genes.

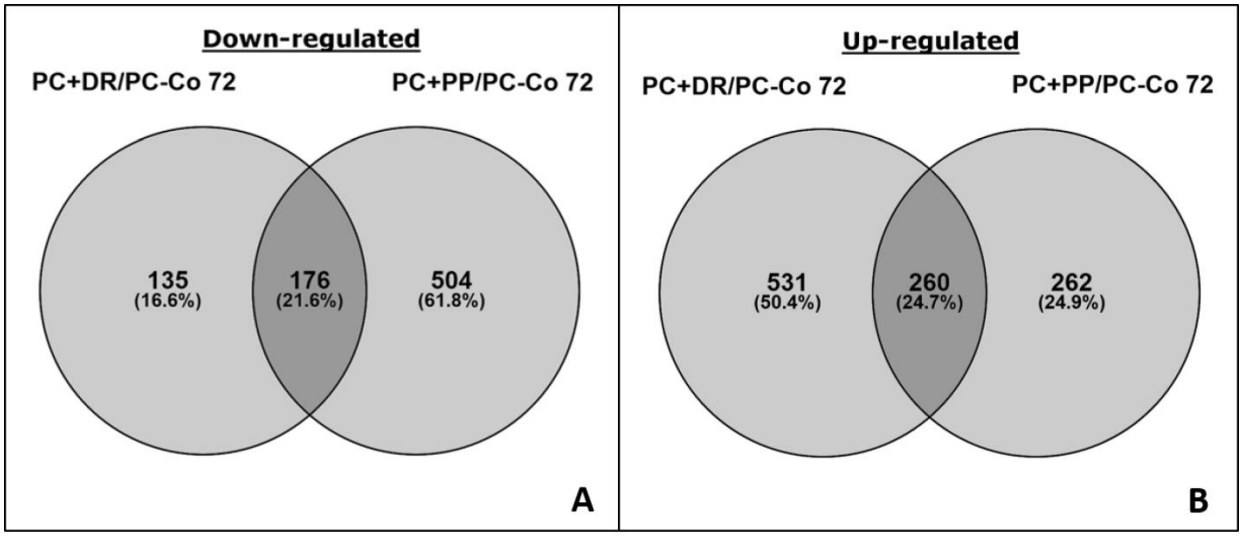


Fig. 8: Venn diagram of the quantitative comparison of the significantly regulated genes in response to inoculations with *D. rosae* and *P. pannosa*. Down-regulated genes are shown in A and up-regulated genes in B.

Table 3: Examples of the most highly up- and down-regulated genes according to their mean log2-fold change. The table is divided into three parts according to the gene expression pattern. Significant changes in expression are marked by asterisks. ↑ indicates that the gene was not expressed in the control treatment, so the extent of the regulation was not quantified. - indicates that the gene was detected neither in the control treatment nor in the inoculated samples.

Feature ID	Blast2GO Sequence Description	Log2 FC PC+DR_24/PC-Co_24	Log2 FC PC+DR_72/PC-Co_72	Log2 FC PC+PP_72/PC-Co_72
Most strongly regulated genes in the rose-black spot interaction				
RC5G0388500	---NA---	↑	↑*	↑
RC3G0383800	calmodulin 11	4.9*	↑*	↑
RC0G0157700	receptor kinase FERONIA	4.5*	↑*	↑
RC6G0485000	type I inositol 1,4,5-trisphosphate 5-phosphatase 11	4.4	↑*	↑
RC4G0172800	major allergen Pru ar 1-like	7.5*	↑*	↑
RC7G0476500	probable disease resistance At5 g66900	-5.0*	-5.6*	-1.3
RC7G0476600	probable disease resistance At5 g66900	-5.7*	-5.5*	-1.6
RC3G0045900	U-box domain-containing 21-like	-4.9	-5.0*	-1.6
RC5G0606100	probable phosphatase 2C 25	-4.0*	-4.1*	-1.4
RC5G0386200	---NA---	-3.7*	-4.0*	-1.1
Most strongly regulated genes in the rose-powdery mildew interaction				
RC1G0220500	NRT1 PTR FAMILY -like	4.0	↑	↑*
RC1G0288300	Gag-Pol poly	4.9*	↑	↑*
RC1G0518300	21 kDa -like	↑	↑	↑*
RC2G0123600	---NA---	-	↑	↑*
RC2G0281500	clavaminate synthase At3g21360	2.0	↑	↑*
RC3G0262800	2OG-Fe(II) oxygenase family oxidoreductase	-0.9	-1.8	-6.0*
RC2G0592100	collagen alpha-2(IV) chain-like	-3.6	-1.0	-5.9*
RC5G0595100	ammonium transporter 1 member 3-like	-1.9	-1.4	-5.5*
RC7G0395600	L-ascorbate oxidase-like	-1.7	-1.8	-5.4*
RC5G0224900	cytochrome P450 CYP749A22-like	0.9	-1.2	-5.1*

Table 3 continued: Examples of the most highly up- and down-regulated genes according to their mean log2-fold change. The table is divided into three parts according to the gene expression pattern. Significant changes in expression are marked by asterisks. ↑ indicates that the gene was not expressed in the control treatment, so the extent of the regulation was not quantified. - indicates that the gene was detected neither in the control treatment nor in the inoculated samples.

Feature ID	Blast2GO Sequence Description	Log2 FC PC+DR_24/PC-Co_24	Log2 FC PC+DR_72/PC-Co_72	Log2 FC PC+PP_72/PC-Co_72
Most strongly regulated genes in the interaction with both pathogens				
RC5G0038900	cysteine-rich receptor kinase 25	4.5*	↑*	↑*
RC2G0684800	probable WRKY transcription factor 75	7.1*	↑*	↑*
RC5G0394900	cytochrome P450 CYP736A12-like	8.1*	↑*	↑*
RC0G0120200	---NA---	3.6	↑*	↑*
RC4G0174500	major allergen Pru ar 1-like	9.5*	10.4*	7.4*
RC2G0104400	21 kDa	-4.9*	-4.8*	-5.4*
RC0G0160900	cytochrome P450 704C1-like	-4.6*	-4.6*	-6.0*
RC2G0523700	7-deoxyloganetin glucosyltransferase-like	-4.2	-4.1*	-5.3*
RC5G0576100	tryptophan synthase alpha chain-like	-5.7*	-5.1*	-4.3*
RC0G0055100	transcription factor bHLH92	-6.5*	-5.5*	-3.8*

Similarities in the responses to both pathogens

The GO enrichment of the genes that were up-regulated in both interaction systems at 72 hpi showed many defense-related GO terms (Supplementary Table 3), indicating a shared defense response to both pathogens.

Many PR genes, the main indicators of a defense response (Bowles, 1990), were strongly up-regulated in the interaction with both pathogens. In particular, many PR10 genes and major allergens (Pru av1, Pru ar1, Mal d 1) belonging to this class of PR-genes (Liu and Ekramoddoullah, 2006) showed very high levels of up-regulation. Of the 39 major allergens found in the genome of 'Old Blush', 29 were highly up-regulated in response to *D. rosae*, and the majority of these were already differentially expressed at 24 hpi. Of these, 15 were also significantly up-regulated in response to *P. pannosa*, and eight of them were up-regulated at 24 hpi. In addition to PR10, the up-regulation of chitinases, which belong to the PR3 and PR4 classes, was observed in both pathosystems. Eleven chitinases were up-regulated in response to *D. rosae*, and some of them were more highly up-regulated at 24 hpi than at 72 hpi. Six of them were also up-regulated in response to *P. pannosa*, and four of these were already up-regulated at 24 hpi. In addition, six genes only annotated as "pathogenesis-related family protein" were up-regulated in response to *D. rosae*, and five were up-regulated in response to *P. pannosa*.

Plant hormones are an important part of the signaling cascade in response to biotic stress (Yang et al., 2015). Different ethylene response factors (ERFs) showed high levels of up- and down-regulation, many with similar reactions, indicating a response of the ethylene pathway. In addition, eight 1-aminocyclopropane-1-carboxylate oxidases (AAC oxidases), which catalyze the last step of ethylene biosynthesis, were up-regulated in response to *D. rosae* at 24 hpi, and five of them were up-regulated at 72 hpi. Some of the AAC oxidase genes also showed a slight level of up-regulation in response to *P. pannosa*, but none of them were significantly differentially expressed, according to the applied criteria.

The GO enrichment analysis indicated the auxin signaling pathway was down-regulated in both pathosystems. In particular, various auxin response genes, such as small auxin up RNAs (SAURs), auxin response factors (ARFs), auxin-binding proteins (ABPs) and other auxin up-regulated AUX/IAA genes were down-regulated.

The network of WRKY transcription factors is involved in many stress responses, and some of them are key factors in the defense response (Phukan et al., 2016). The inoculations with both fungal pathogens led to the up-regulation of numerous WRKYs, specifically WRKY27, WRKY31, WRKY47, WRKY48, WRKY61and two WRKY71s. Of particular interest are two WRKY75 genes, one was exclusively expressed in the samples inoculated with both pathogens at 72 hpi, while no expression was detected in the control samples (Table 4). The second WRKY75 gene was strongly up-regulated in both interaction systems, even if it did not reach the significance threshold in the interaction with P. pannosa (FDR adjusted p-value 0.063).

Genes specifically regulated in reaction to *P. pannosa*

Surprisingly, the GO enrichment (Supplementary Table 3) of the genes exclusively up-regulated in response to *P. pannosa* infection did not result in any overrepresented GO terms. Thus, no PR genes or other genes related to a defense response specific to the interaction of roses to powdery mildew could be identified in this set of genes. However, the enrichment analysis of exclusively down-regulated genes indicated that processes related to photosynthesis and cell wall organization were affected.

The majority of the significantly down-regulated genes associated with photosynthesis code for chlorophyll a/b binding proteins (*cab*), structural components of photosystem I (subunit l, K, N) and a PsbP gene of photosystem II. In addition, four genes of different steps in chlorophyll biosynthesis were down-regulated, namely, glutamyl-tRNA reductase, a subunit of magnesium-chelatase, geranylgeranyl diphosphate reductase and a chlorophyllase (Tripathy and Pattanayak, 2012).

Several genes related to cell-wall organization were only down-regulated in response to *P. pannosa*, including genes encoding for six expansins, three trichome birefringence-like proteins (TBL), two fasciclin-like arabinogalactan (FLA) proteins and also two COBRA-like genes and two walls are thin1 (WAT1) genes. Additionally, one cellulose synthase 4 gene, three genes encoding for pectin lyases, two xyloglucan endotransglucosylase genes and some genes of the lignin biosynthesis pathway were down-regulated.

Genes specifically regulated in response to *D. rosae*

Among the genes specifically regulated in response to the interaction with *D. rosae*, PR5 genes (thaumatin) were highly up-regulated at 24 and 72 hpi, and a PR1 gene was among the most highly up-regulated genes.

Furthermore, the GO enrichment analysis (Supplementary Table 3) indicated that genes of the secondary metabolism were up-regulated in response to *D. rosae*. In particular, the phenylpropanoid and the flavonoid pathways seemed to be affected. Fig. 9 displays the main steps of the flavonoid synthesis pathway and the significantly regulated genes. It can be seen that, in response to *D. rosae*, genes in almost all the steps were up-regulated. In particular, the genes encoding for chalcone isomerases (CHI), dihydroflavonol 4-reductase (DFR) and flavonol synthase (FLS) were strongly up-regulated at both time points, indicating that the synthesis of flavonols was induced. In addition, the pathway leading to the synthesis of proanthocyanidins (DFR, LDOX, ANR) was up-regulated. The only down-regulated genes in this interaction were two genes encoding UDPG-flavonoid glucosyl transferase (UFGT). This is a key enzyme in the synthesis of anthocyanins. In contrast to the reaction of PC to *D. rosae*, the reaction of PC to *P. pannosa* displayed a different gene regulation pattern. There were significantly up-regulated genes, such as one of the flavanone 3-hydroxylase (F3H) genes, one FLS gene and one leucoanthocyanidin dioxygenase (LDOX) gene. In particular, enzymes acting in the first steps of the flavonoid synthesis pathway (PAL, 4CL, CHS and CHI) seemed to remain unaffected or were even down-regulated. In addition, the anthocyanidin reductase (ANR) gene leading to the synthesis of proanthocyanidins was significantly down-regulated. Furthermore, the flavonol synthesis pathway was not as strongly induced in the interaction to *P. pannosa* as it was in the response to *D. rosae*; only one of the three FLS genes was up-regulated.

Many genes encoding key enzymes in the phenylpropanoid pathway were up-regulated in most of the cases at both time points, 24 hpi and 72 hpi, in response to *D. rosae*, with *p*-coumarate 3-hydroxylase (C3H), caffeic acid *O*-methyltransferase (COMT), caffeoyl-CoA *O*-methyltransferase (CCoAOMT) and cinnamoyl-CoA reductase (CCR) showing the strongest up-regulation (Fig. 10). Of particular interest are CCoAOMT and COMT because they are involved in the synthesis of G- and S-lignin, which are the main components of lignin in dicots (Vanholme et al., 2010). In

addition to the biosynthesis genes, several peroxidases and laccases, which are involved in the cross-linking of lignin monomers, were up-regulated. This pathway seemed to be less affected in response to *P. pannosa* than in response to *D. rosae*. One CCR gene and a COMT gene showed higher levels of up-regulation, but some key factors were also significantly down-regulated. Notably, two genes encoding for cinnamyl alcohol dehydrogenase (CAD) were significantly down-regulated. This enzyme catalyzes one of the later steps in the synthesis of all three types of lignin monomers.

Remarkably, genes responding to the salicylic acid (SA)-mediated signaling pathway were only up-regulated in the samples inoculated with *D. rosae*. Various response genes were significantly up-regulated, including typical marker genes, such as phytoalexin deficient4 **(**PAD4), PR1 and PR5s (Derksen et al., 2013). The up-regulation of the PR1 gene was also observed in the qPCR validation. In addition, various senescence-associated carboxylesterase 101 (SAGs 101) genes were up-regulated after 72 hpi. An EDS1 gene was also up-regulated at this time point, even if its regulation did not completely meet the stringent criteria applied in this study.

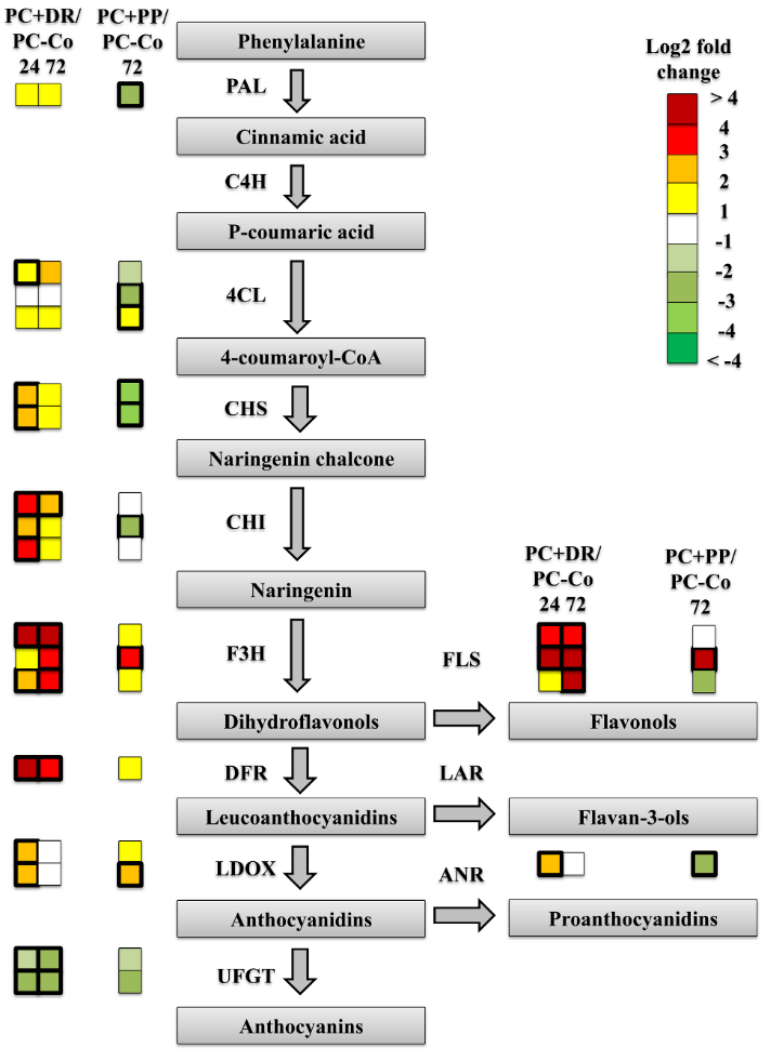


Fig. 9: Overview of the regulation of the flavonoid biosynthesis pathway. The main steps in the flavonoid biosynthesis pathway and the regulation of genes in the pathway in response to the two pathogens are displayed. Each quadrat represents a gene. Significant changes in gene expression are indicated by thicker lines. Enzyme names are abbreviated as follows: PAL: phenylalanine ammonia-lyase, C4H cinnamate-4-hydroxylase, 4CL 4-coumaroyl-CoA-ligase, CHS chalcone synthase, CHI chalcone isomerase, F3H flavanone 3-hydroxylase, DFR dihydroflavonol 4-reductase, LDOX leucoanthocyanidin dioxygenase, UFGT UDPG-flavonoid glucosyl transferase, FLS flavonol synthase, LAR leucoanthocyanidin reductase, ANR anthocyanidin reductase.

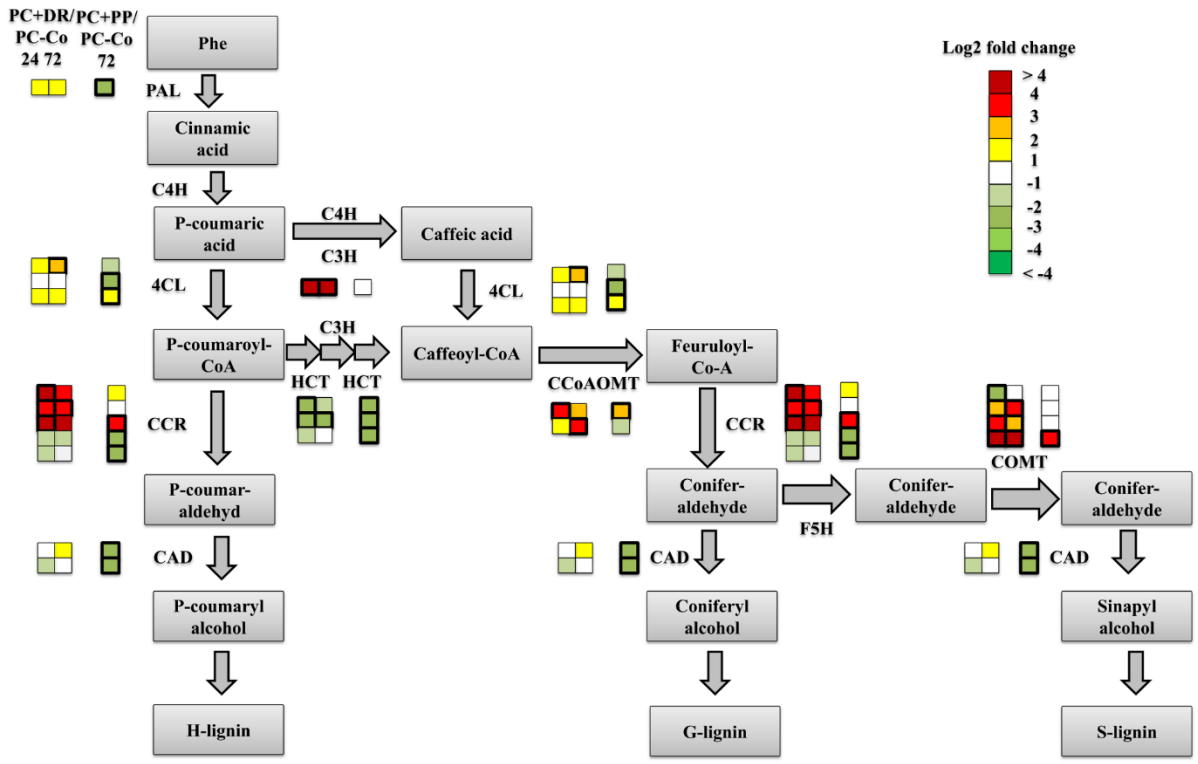


Fig. 10: Overview of the regulation of the phenylpropanoid biosynthesis pathway. The main steps in the phenylpropanoid biosynthesis pathway and the regulation of genes in the pathway in response to the two pathogens are displayed, significant changes in gene expression are marked with thicker lines. Enzyme names are abbreviated as follows: PAL: phenylalanine ammonia-lyase, C4H: cinnamate-4-hydroxylase, 4CL: 4-coumaroyl-CoA-ligase, HCT: hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase, C3H: *p*-coumarate 3-hydroxylase, CCoAOMT: caffeoyl-CoA *O*-methyltransferase, CCR: cinnamoyl-CoA reductase, F5H: ferulate 5-hydroxylase, COMT: caffeic acid *O*-methyltransferase, and CAD: cinnamyl alcohol dehydrogenase.

Discussion

Application of the MACE technology

We analyzed changes in the rose leaf transcriptome during the early stages of compatible interactions with two pathogens, the hemibiotrophic fungus *D. rosae* and the obligate biotrophic fungus *P. pannosa*, using the MACE technique. The sequencing (6-30 million reads per library) and mapping (76.3-92.9%) results are comparable to those of other studies that applied this technique, e.g., Hradilová et al. (2017), applied the MACE technique to pea, with an output of 8 to 15 million reads per library, and 12.3– 21.7 million reads per library were generated from the RNA of apple roots (Weiß et al., 2017). A noteworthy difference between MACE and

conventional RNA-seq technology is that, in the former, only one read is produced for each cDNA molecule. Although the read coverage between RNA-seq and MACE is difficult to compare, the 6-30 million tags obtained in the present study reflect a much larger number of tags in RNA-seq analysis, where each cDNA might be represented by several reads. The 44481 predicted transcripts of the rose genome (Hibrand et al., 2018) have an average total length of 2414 base pairs (data not shown), which is between the previously reported mean transcript length for *Arabidopsis*, with a length of 2411 bp, and rice, with a length of 3435 bp (Ren et al., 2006). In the RNA-seq-approach, each cDNA molecule results in several sequenced reads, depending on the sequencing technique used and the transcript length, whereas in the MACE approach, only one tag per cDNA-molecule is generated. It is conservative to assume that the average number of tags per transcript in an RNA-seq analysis is larger than 10. Therefore, the coverage obtained with our MACE analyses exceeds that of most RNA-Seq approaches that have individual library sizes between 30 and 50 million reads.

The majority of significantly differentially expressed genes indicated by the MACE technique could be validated with RT-qPCR. Taking into account that all three biological repeat experiments in the MACE analysis were derived from independent inoculation experiments and that the validation included three additional independent inoculation experiments, which introduced more biological and technical variability then conventional repeats, the correlation coefficient of 0.82 is surprisingly high. However, this might be due to the stringent conditions we applied for the identification of differentially expressed genes. Nevertheless, the MACE data still contains effects, which might have been caused by outliers in one biological repeat. Thus, for a meaningful interpretation of the data, all repeats have to be considered separately, not just the mean fold changes. The clustering of the samples from PC+DR 24, PC+DR 72 and PC+PP 24 into one distinct cluster (Fig. 2) indicates that there is a clear effect of the inoculation with both pathogens. The fact that all samples at 0 hpi are separated from the samples at the other time points indicates that the experiment procedure caused additional stress also visible in the 24 hpi and 72 hpi control samples. The subclusters of the data derived from the third biological repetition can be explained by the fact that these samples were generated 12 months later than those of the other two repeat experiments. In addition, the reads of the third biological repeat are longer than those of the other two repeats (Table 1), due to

changes in the sequencing procedure at GenXPro. Nevertheless, the effect of the inoculation was still stronger than that of the biological repetition, as indicated by the three main clusters.

Comparison of the response of roses to *D. rosae* and *P. pannosa*

The comparison of the transcriptomic changes of the susceptible rose variety PC in response to the pathogens *D. rosae* and *P. pannosa* indicates that there are two defense responses. One seems to be a more general response and is similar in both pathosystems; it is characterized by the up-regulation of PR10 genes and chitinases. The second type of reaction is specific to *D. rosae* and includes the up-regulation of additional PR genes and secondary metabolism genes.

The response to both pathogens

The up-regulation of chitinases and PR10 is often reported as a response to biotic stress. In particular, PR10 genes seem to play a central role in the defense response of Rosaceae. Studies of the defense response of the Rosaceae model plant *Fragaria* infected by pathogens with different lifestyles, such as the necrotrophic fungus *Botrytis cinerea* (González et al., 2013), the obligate biotrophic powdery mildew fungus *Podospaera aphanis* (Jambagi and Dunwell, 2015) and the hemibiotrophic oomycete *Phytophthora cactorum* (Toljamo et al., 2016), showed a strong up-regulation of PR10 genes or major allergens belonging to this family. Studies on apple (*Malus*) infected with *Venturia inaequalis* (apple scab) (Poupard et al., 2003, Cova et al., 2017) and treated with fungal elicitors (Pühringer et al., 2000) also showed an up-regulation of PR10 genes. The biological functions of the PR10 gene family are not completely known, but, among other activities, an antifungal activity has been described (Flores et al., 2002). The role of chitinases in the defense response of plants has been analyzed extensively in different systems; they block hyphal growth, trigger other defense mechanisms through the release of elicitors and play a role in the so-called “hypersensitive response” (Grover, 2012).

Another remarkable aspect of the up-regulation of PR10 genes and chitinases in response to infection by *D. rosae* and *P. pannosa* is the early occurrence in the

infection process at 24 hpi. This result is even more notable because very few other changes in the leaf transcriptome were detected in response to *P. pannosa* at 24 hpi. One of the reasons for the low number of significantly regulated genes at this time point could be that the inoculation density was lower for *P. pannosa* than for *D. rosae*. The fact that many PR10 genes and chitinases had much higher expression levels in response to *D. rosae* than to *P. pannosa* at 72 hpi indicates that the inoculation density might have influenced the results. However, this quick and conserved response, as well as the fact that we only analyzed the compatible interactions, might indicate that it was a PTI reaction elicited by chitin or the penetration of the cuticle, highlighting the importance for general defense reactions.

Many WRKY transcription factors were up-regulated in response to both pathogens, indicating that they might play an important role in activating similar response processes to pathogen attacks. The up-regulation of various WRKYs in response to *Marssonina brunnea*, which is closely related to *D. rosae*, was also observed in a transcriptomic study of poplar (*Populus*). The majority of the WRKYs were also up-regulated in response to other stress factors (Jiang et al., 2014). However, in poplar, a WRKY75 gene reacted only to biotic stress, hormone treatments and salt stress. WRKY75 also seems of particular interest in the response of roses to *D. rosae* and *P. pannosa* because it was exclusively expressed in the inoculated samples. The importance of WRKY75 for the basal immune response and resistance reaction has also been shown for the *Fragaria* homologue of WRKY75 and the *Arabidopsis* WRKY75 (Encinas-Villarejo et al., 2009). Furthermore, a meta-analysis of *Malus* infected with different pathogens showed an up-regulation of WRKY75 in the majority of the interactions (Balan et al., 2018). Additionally, WRKY71, which was also up-regulated in response to both pathogens, has been reported as a positive regulator in the defense response of rice (Liu et al., 2007).

The down-regulation of the auxin-mediated pathway in both pathosystems might indicate an inactivation of developmental processes and plant growth. It has frequently been shown that plants reduce these processes during pathogen infection to divert their limited resources to the defense response (Kazan and Manners, 2009).

The role of ethylene as a defense regulator has been well-established, mainly due to its function as a signaling molecule in the defense against necrotrophic pathogens in combination with jasmonic acid. Nevertheless, there have also been reports that

ethylene is a major factor controlling defense responses, including those against biotrophic pathogens (Yang et al., 2015). Therefore, a role of ethylene as a signaling molecule in the presented pathosystems is possible.

Changes in the transcriptome specific to the interaction with *P. pannosa*

Surprisingly, no specific group of genes related to common defense functions were found to be up-regulated specifically in the reaction to *P. pannosa*. As mentioned above, this may be due to the low *P. pannosa* inoculation density and the fact that the infection zone was restricted to the epidermis. Thus, genes with low expression levels might remain below the detection level. However, it seems that genes involved in the light reaction of photosynthesis were down-regulated more strongly in response to *P. pannosa* than in response to *D. rosae,* which is surprising, considering that *P. pannosa* only infects epidermal cells, which are not photosynthetically active, whereas *D. rosae* infects the cell layers below the epidermis. The down-regulation of photosynthesis genes is often observed in plant-pathogen interactions (e.g., Cremer et al., 2013, Milli et al., 2012, Balan et al., 2018). This reaction is explained by a pathogen-induced source-sink transition of infected leaf tissue often accompanied by an up-regulation of extracellular invertases and respiration genes (Berger et al., 2007). However, this is not supported by our data, except for the observation that some sugar transporters were strongly up-regulated. Further studies are needed to determine the influence of *P. pannosa* on photosynthesis and sugar metabolism.

Another process strongly regulated in response to *P. pannosa* was the cell wall modification process. Different factors, such as expansins, TBLs or FLAs, were down-regulated in response to the pathogen. They are known to be involved in modification processes such as cell wall loosening, plant cell expansion and secondary cell wall formation (Cosgrove, 2000, Schindelman et al., 2001, Johnson et al., 2003, Van Sandt et al., 2007, Bischoff et al., 2010). Of particular interest might be WAT1 genes. Two WAT1 gene homologues were expressed in our samples, and one was down-regulated, with a generally low expression level in both interaction systems, while the other, which had a much higher expression level, was exclusively down-regulated in response to *P. pannosa* at 72 hpi. WAT1 was identified as an essential factor for secondary cell wall formation by loss-of-function mutations in

Arabidopsis (Ranocha et al., 2010). In addition, *wat1* mutants showed a dwarf phenotype and down-regulation of almost all genes in the lignin biosynthesis pathway. The down-regulation of the highly expressed WAT1 gene exclusively observed in response to *P. pannosa* might be an important factor leading to the observed differences in cell wall formation and the lignin pathway, and it might be one of the ways the fungus suppresses the plant defense.

Changes in the transcriptome specific to the interaction with *D. rosae*

The transcriptomic changes specific to the interaction with *D. rosae* include the up-regulation of many genes encoding enzymes in the lignin biosynthesis pathway and the flavonoid pathway. The synthesis of lignin or lignin-like phenolic polymers is a common phenomenon in response to pathogens. For instance, Chinese cabbage plants infected with the necrotrophic bacterium *Erwinia carotovora* accumulated high levels of lignin monomers, and 12 genes involved in lignin biosynthesis were up-regulated (Zhang et al., 2007). Infections with the fungus *Sclerotinia sclerotium* led to the lignification of *Camelina sativa* cells and were correlated with an up-regulation of CCR genes (Eynck et al., 2012). Additionally, the defense response of the model plant *Arabidopsis thaliana* infected with *Pseudomonas syringae* was accompanied by the up-regulation of CAD genes. Loss-of-function mutations of these genes resulted in plants with a reduced resistance to the pathogen (Tronchet et al., 2010). Lignin has different functions in the plant defense. As part of the secondary cell wall, it provides a physical barrier against the entry of the pathogen. Simultaneously, it prevents the spread of pathogen toxins and enzymes into neighboring cells and the transfer of water and nutrients from the host to the pathogen. In addition, cell wall components can have signaling functions in defense (Miedes et al., 2014). As mentioned above, WAT1 is an essential factor in the formation of the secondary cell wall and influences the lignin biosynthesis pathway (Ranocha et al., 2010). One of the two expressed homologues of WAT1, which had much higher expression values than the other homolog, was down-regulated in response to *P. pannosa* and was unaffected by *D. rosae*. Therefore, it could be one of the factors leading to the different reactions of this pathway in response to the two pathogens.

In contrast to the lignin biosynthesis pathway, genes encoding for almost all enzymes of the flavonoid pathway were up-regulated specifically in response to *D. rosae*. Products of the flavonoid pathway have various functions in plants, e.g., they act as chemical messengers that interact with insects and microbes, function as pigments to attract pollinators or protect against UV light. More importantly, many phytoalexins are synthesized by this pathway (Piasecka et al., 2015, Falcone Ferreyra et al., 2012). In *Fragaria*, catechin and catechin-derived proanthocyanidins have been shown to be involved in the defense response to *Alternaria alteranta* and *B. cinerea* in infected leaves and fruits, respectively (Yamamoto et al., 2000, Puhl and Treutter, 2008). Additionally, a recent transcriptomic study of roots infected with *P. cactorum* showed an up-regulation of the flavonoid pathway genes leading to these products (Toljamo et al., 2016). Interestingly, in our data, FLS genes were highly up-regulated, indicating that the synthesis of flavonols is more important in the response to *D. rosae* than in the response to *P. pannosa*. In plants, flavonols function as antioxidants during high light conditions and as detoxifying agents against ROS (reactive oxygen species), which might explain their role in the plant defense (Pollastri and Tattini, 2011). Metabolic analyses focusing on this group of metabolites might be needed to confirm the induction of this pathway in the rose-*D. rosae* interaction.

Different SA response genes were exclusively up-regulated in response to *D. rosae*. PAD4, EDS1 and SAG101 cooperate in stimulating the production of SA and are essential for SA-mediated defense responses (Zhou et al., 1998, Feys et al., 2005, Rietz et al., 2011), which are typically characterized by the up-regulation of PR1, PR2 and PR5 genes. PR1 and different thaumatins (PR5) were also exclusively up-regulated in response to *D. rosae,* except for one PR5 paralogue that was also up-regulated in response to *P. pannosa*. The up-regulated PR1 gene was included in the expression analysis for the validation of the MACE data with the BioMark qPCR system. This analysis confirmed the results of the transcriptomic data and showed a strong up-regulation (22-fold at 24 hpi, 10-fold at 72 hpi) of the PR1 gene in response to *D. rosae.* Additionally, our analyses indicated significant differential expression of the same gene in response to *P. pannosa*, with a 4-fold down-regulation at 24 hpi and a 3-fold down-regulation at 72 hpi (Supplementary Table 1). This response is surprising, because SA is known to play a key role in activating the defense response to biotrophic and hemibiotrophic pathogens (van Loon et al., 2006, Derksen et al.,

2013). This could be an indication that the additional *D. rosae*-specific stress responses are regulated by the SA-mediated signaling pathway, which is either suppressed during the interaction with *P. pannosa* or not up-regulated in the *P. pannosa* pathosystem because *D. rosae* is recognized differently. This could involve, for example, a partial ETI reaction, where specific effectors of *D. rosae* are detected. We know from previous experiments that there are more than 20 genes in the rose genome that are similar to the resistance-mediating *muRdr1A* gene of the Rdr1-locus (Terefe-Ayana et al., 2011) but do not mediate a full resistance. Nevertheless, they might function in activating a partial ETI reaction, leading to the observed differences in the reaction of roses to the pathogens.

Conclusions

With this first analysis of the rose defense transcriptome we could show contrasting responses of the host depending on the different lifestyles of two fungal pathogens, the hemibiotrophic *Diplocarpon rosae* and the biotrophic *Podosphaera pannosa*. Besides a common response to both pathogens, characterized by an up-regulation of PR10 genes and chitinases, processes like photosynthesis and cell wall modification were mainly down-regulated in response to *P. pannosa*, whereas the secondary metabolism in form of the phenylpropanoid and flavonoid pathway was mainly up-regulated in response to *D. rosae*. Surprisingly, PR1 and the SA-pathway were exclusively up-regulated in response to the hemibiotrophic *D. rosae* and not as expected in the biotrophic interaction with *P. pannosa*. Some of the described mechanisms were already observed in other Rosaceae species, indicating the importance for the defense response and the conservation of these processes. This information is an important first step in understanding the response to both rose pathogens and revealed many processes, which need to be analyzed in more detail.

Author contributions statement

The contribution of the authors to this study is as follows:

EN was involved in conceptualizing the project, performing and evaluating the experiments, interpreting the data and writing and reviewing the article.

HD was involved in performing and evaluating the qPCR experiments, evaluating and interpreting the MACE data and writing and reviewing the article.

HK was involved conceptualizing the project, performing the experiments for the generation of the MACE data and writing and reviewing the article.

ML was involved in writing and reviewing the article.

IM was involved in performing and evaluating the qPCR experiments, evaluating and interpreting the MACE data and writing and reviewing the article.

TD was involved in conceptualizing and supervising the project, interpreting the data and writing and reviewing the article.

Funding disclosure

This work was funded by the DFG GRK 1798 "Signaling at the Plant-Soil Interface" (GRK1798/1).

Acknowledgments

We want to give special thanks to the Institut für Rebenzüchtung of the Julius Kühn-Institut (JKI) in Siebeldingen. In particular, we thank the group of Prof. Dr. Eva Zyprian for the support in performing high-throughput qPCR with the Biomark system. Additionally, we thank the DFG for the funding of this project in the background of the GRK 1798.

Conflict of Interest Statement

The authors declare no conflict of interest.

Supplementary Material

Supplementary Table 1: MACE validation

Supplementary Table 2: MACE data, PC+PP vs. PC-Co

Supplementary Table 3: GO Enrichment

Supplementary Table 4: MACE data, PC+DR vs. PC-Co

References

Aronescu, A (1934). *Diplocarpon Rosae*: From spore germination to haustorium formation. Bull Torr Bot Club 61, 291-329. doi: 10.2307/2480882

Ashburner, M, Ball, C A, Blake, J A, Botstein, D, Butler, H, Cherry, J M, et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25, 25–29. doi: 10.1038/75556

Balan, B, Marra, F P, Caruso, T, and Martinelli, F (2018). Transcriptomic responses to biotic stresses in Malus x domestica: a meta-analysis study. Sci Rep. 8(1):1970. doi: 10.1038/s41598-018-19348-4

Bent, A F, and Mackey, D (2007). Elicitors, effectors, and R genes: The new paradigm and a lifetime supply of questions. Annu Rev Phytopathol 45, 399–436. doi: 10.1146/annurev.phyto.45.062806.094427

Berger, S, Sinha, A K, and Roitsch, T (2007). Plant physiology meets phytopathology: plant primary metabolism and plant-pathogen interactions. J Exp Bot 58, 4019–4026. doi: 10.1093/jxb/erm298

Bischoff, V, Nita, S, Neumetzler, L, Schindelasch, D, Urbain, A., Eshed, R., et al. (2010). TRICHOME BIREFRINGENCE and its homolog AT5G01360 encode plant-specific DUF231 proteins required for cellulose biosynthesis in Arabidopsis. Plant Physiol 153(2), 590–602. doi: 10.1104/pp.110.153320

Bowles, D J (1990). Defense-related proteins in higher plants. Annu Rev Biochem 59, 873–907. doi: 10.1146/annurev.bi.59.070190.004301

Conesa, A, Gotz, S, Garcia-Gomez, J M, Terol, J, Talon, M, and Robles, M (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21, 3674–3676. doi: 10.1093/bioinformatics/bti610

Cosgrove, D J (2000). Loosening of plant cell walls by expansins. Nature 407, 321–326. doi: 10.1038/35030000

Cova, V, Paris, R, Toller, C, Patocchi, A, Velasco, R, and Komjanc, M (2017). Apple genes involved in the response to *Venturia inaequalis* and salicylic acid treatment. Sci Hortic 226, 157–172. doi: 10.1016/j.scienta.2017.08.043

Coyier, D L (1983). Control of rose powdery mildew in the greenhouse and field. Plant Disease 67, 919–923.

Cremer, K de, Mathys, J, Vos, C, Froenicke, L, Michelmore, R W, Cammue, Bruno P A, et al. (2013). RNAseq-based transcriptome analysis of *Lactuca sativa* infected by the fungal necrotroph *Botrytis cinerea*. Plant Cell Environ 36, 1992–2007. doi: 10.1111/pce.12106

Davies, D R (1980). Rapid propagation of roses in vitro. Sci Hortic 13, 385–389. doi: 10.1016/0304-4238(80)90097-7

Debener, T, and Byrne, D H (2014). Disease resistance breeding in rose: Current status and potential of biotechnological tools. Plant Sci 228, 107-117. doi: 10.1016/j.plantsci.2014.04.005

Debener, T, Drewes-Alvarez, R, and Rockstroh, K (1998). Identification of five physiological races of blackspot, *Diplocarpon rosae*, Wolf on roses. Plant Breeding 1998, 267–270. doi: 10.1111/j.1439-0523.1998.tb01937.x

Derksen, H, Rampitsch, C, and Daayf, F (2013). Signaling cross-talk in plant disease resistance. Plant Sci 207, 79–87. doi: 10.1016/j.plantsci.2013.03.004

Dohm, A, Ludwig, C, Nehring, K, and Debener, T (2001) Somatic embryogenesis in roses. Acta Hortic 547, 341–347. doi: 10.17660/ActaHortic.2001.547.40

Dugo, M L, Satovic, Z, Millán, T, Cubero, J I, Rubiales, D., Cabrera, A., et al. (2005). Genetic mapping of QTLs controlling horticultural traits in diploid roses. Theor Appl Genet 111, 511–520. doi: 10.1007/s00122-005-2042-4

Encinas-Villarejo, S, Maldonado, A M, Amil-Ruiz, F, de los Santos, B, Romero, F, Pliego-Alfaro, F, et al. (2009). Evidence for a positive regulatory role of strawberry (*Fragaria x ananassa*) Fa WRKY1 and Arabidopsis At WRKY75 proteins in resistance. J Exp *Bot.* 60, 3043–3065. doi: 10.1093/jxb/erp152

Eynck, C, Seguin-Swartz, G, Clarke, W E, and Parkin, I A (2012). Monolignol biosynthesis is associated with resistance to *Sclerotinia sclerotiorum* in *Camelina sativa*. Mol Plant Pathol 13, 887–899. doi: 10.1111/j.1364-3703.2012.00798.x

Falcone Ferreyra, M L, Rius, S P, and Casati, P (2012). Flavonoids: biosynthesis, biological functions, and biotechnological applications. Front Plant Sci 3, 222. doi: 10.3389/fpls.2012.00222

Feys, B J, Wiermer, M, Bhat, R A, Moisan, L J, Medina-Escobar, N, Neu, C, et al. (2005). Arabidopsis SENESCENCE-ASSOCIATED GENE101 stabilizes and signals within an ENHANCED DISEASE SUSCEPTIBILITY1 complex in plant innate immunity. Plant Cell 17, 2601–2613. doi: 10.1105/tpc.105.033910

Flores, T, Alape-Giron, A, Flores-Diaz, M, and Flores, H E (2002). Ocatin. A novel tuber storage protein from the Andean tuber crop oca with antibacterial and antifungal activities. Plant Physiol 128, 1291–1302. doi: 10.1104/pp.010541

Gachomo, E W, Dehne, H-W, and Steiner, U (2006). Microscopic evidence for the hemibiotrophic nature of Diplocarpon rosae, cause of black spot disease of rose. Physiol Mol Plant Pathol 69, 86–92. doi: 10.1016/j.pmpp.2007.02.002

González, G, Fuentes, L, Moya-León, M A, Sandoval, C, and Herrera, R (2013). Characterization of two PR genes from *Fragaria chiloensis* in response to *Botrytis cinerea* infection: A comparison with *Fragaria x ananassa*. Physiol Mol Plant Pathol 82, 73–80. doi: 10.1016/j.pmpp.2013.02.001

Grover, A (2012). Plant Chitinases: Genetic diversity and physiological roles. Crit Rev Plant Sci 31, 57–73. doi: 10.1080/07352689.2011.616043

Hattendorf, A, Linde, M, Mattiesch, L, Debener, T, and Kaufmann, H (2004). Analysis of rose resistance genes and their localisation in the rose genome. Acta Hortic 2004, 123–130. doi: 10.17660/ActaHortic.2004.651.14

Horbach, R, Navarro-Quesada, A R, Knogge, W, and Deising, H B (2011). When and how to kill a plant cell: Infection strategies of plant pathogenic fungi. J Plant Physiol 168, 51–62. doi: 10.1016/j.jplph.2010.06.014

Hosseini Moghaddam, H, Leus, L, Riek, J de, van Huylenbroeck, J, and van Bockstaele, E (2012). Construction of a genetic linkage map with SSR, AFLP and

morphological markers to locate QTLs controlling pathotype-specific powdery mildew resistance in diploid roses. Euphytica 184, 413–427. doi: 10.1007/s10681-011-0616-6

Hradilová, I, Trněný, O, Válková, M, Cechová, M, Janská, A, Prokešová, L, et al. (2017). A combined comparative transcriptomic, metabolomic, and anatomical analyses of two key domestication traits: Pod dehiscence and seed dormancy in pea (Pisum sp.). Front Plant Sci 8, 542. doi: 10.3389/fpls.2017.00542

Jambagi, S, and Dunwell, J M (2015). Global transcriptome analysis and identification of differentially expressed genes after infection of *Fragaria vesca* with powdery mildew (*Podosphaera aphanis*). Transcriptomics 03. doi: 10.4172/2329-8936.1000106

Jiang, Y, Duan, Y, Yin, J, Ye, S, Zhu, J, Zhang, F, et al. (2014). Genome-wide identification and characterization of the Populus WRKY transcription factor family and analysis of their expression in response to biotic and abiotic stresses. J Exp Bot 65, 6629–6644. doi: 10.1093/jxb/eru381

Johnson, K L, Jones, B J, Bacic, A, and Schultz, C J (2003). The fasciclin-like arabinogalactan proteins of arabidopsis. A multigene family of putative cell adhesion molecules. Plant Physiol 133, 1911–1925. doi: 10.1104/pp.103.031237

Jones, J D, and Dangl, J L (2006). The plant immune system. Nature 444, 323–329. doi: 10.1038/nature05286

Kahl, G, Molina, C, Rotter, B, Juengling, R, Frank, A, Krezdorn, N, et al. (2012). Reduced representation sequencing of plant stress transcriptomes. J Plant Biochem Biotechnol 21, 119-127. doi: 10.1007/s13562-012-0129-y

Kaufmann, H, Mattiesch, L, Lörz, H, and Debener, T (2003). Construction of a BAC library of *Rosa rugosa* Thunb. and assembly of a contig spanning Rdr1, a gene that confers resistance to blackspot. Mol Genet Genomics 268, 666–674. doi: 10.1007/s00438-002-0784-0

Kaufmann, H, Qiu, X, Wehmeyer, J, and Debener, T (2012). Isolation, molecular characterization, and mapping of four rose *MLO* orthologs. Front Plant Sci 3:244. doi: 10.3389/fpls.2012.00244

Kazan, K, and Manners, J M (2009). Linking development to defense: auxin in plant–pathogen interactions. Trends Plant Sci 14, 373–382. doi: 10.1016/j.tplants.2009.04.005

Klie, M, and Debener, T (2011). Identification of superior reference genes for data normalisation of expression studies via quantitative PCR in hybrid roses (*Rosa hybrida*). BMC Res Notes 4, 518. doi: 10.1186/1756-0500-4-518

Leus, L, Dewitte, A, van Huylenbroeck, J, Vanhoutte, N, van Bockstaele, E, and Hoefte, M (2006). *Podosphaera pannosa* (syn. *Sphaerotheca pannosa*) on Rosa and Prunus spp. Characterization of pathotypes by differential plant reactions and ITS sequences. J Phytopathol 154, 23–28. doi: 10.1111/j.1439-0434.2005.01053.x

Linde, M, and Debener, T (2003). Isolation and identification of eight races of powdery mildew of roses (*Podosphaera pannosa*) (Wallr. Fr.) de Bary and the genetic analysis of the resistance gene Rpp1. Theor Appl Genet 107, 256–262. doi: 10.1007/s00122-003-1240-1

Linde, M, Hattendorf, A, Kaufmann, H, and Debener, T (2006). Powdery mildew resistance in roses: QTL mapping in different environments using selective genotyping. Theor Appl Genet 113, 1081–1092. doi: 10.1007/s00122-006-0367-2

Linde, M, Mattiesch, L, and Debener, T (2004). Rpp1, a dominant gene providing race-specific resistance to rose powdery mildew (*Podosphaera pannosa*): molecular mapping, SCAR development and confirmation of disease resistance data. Theor Appl Genet 109, 1261–1266. doi: 10.1007/s00122-004-1735-4

Linde, M, Shishkoff, N (2003). “Fungi: Powdery mildew,” in *Encyclopedia of Rose Science.*, eds. A Roberts, T Debener, and S Gudin (Oxford: Elsevier, Academic Press), 158–165.

Liu, J-J, and Ekramoddoullah, A K M (2006). The family 10 of plant pathogenesis-related proteins: Their structure, regulation, and function in response to biotic and abiotic stresses. *Physiol. Mol.* Plant Pathol 68, 3–13. doi: 10.1016/j.pmpp.2006.06.004

Liu, X, Bai, X, Wang, X, and Chu, C (2007). OsWRKY71, a rice transcription factor, is involved in rice defense response. J Plant Physiol 164, 969–979. doi: 10.1016/j.jplph.2006.07.006

Lohse, M, Nagel, A, Herter, T, May, P, Schroda, M, Zrenner, R, et al. (2014). Mercator: a fast and simple web server for genome scale functional annotation of plant sequence data. Plant Cell Environ 37, 1250–1258. doi: 10.1111/pce.12231

Menz, I, Straube, J, Linde, M, and Debener, T (2017). The TNL gene Rdr1 confers broad-spectrum resistance to *Diplocarpon rosae*. Mol Plant Pathol doi: 10.1111/mpp.12589

Miedes, E, Vanholme, R, Boerjan, W, and Molina, A (2014). The role of the secondary cell wall in plant resistance to pathogens. Front Plant Sci 5, 358. doi: 10.3389/fpls.2014.00358

Milli, A, Cecconi, D, Bortesi, L, Persi, A, Rinalducci, S, Zamboni, A, et al. (2012). Proteomic analysis of the compatible interaction between *Vitis vinifera* and *Plasmopara viticola*. J Proteom 75, 1284–1302. doi: 10.1016/j.jprot.2011.11.006

Neu, E, Featherston, J, Rees, J, and Debener, T (2017). A draft genome sequence of the rose black spot fungus *Diplocarpon rosae* reveals a high degree of genome duplication. PloS One 12. doi: 10.1371/journal.pone.0185310

Phukan, U J, Jeena, G S, and Shukla, R K (2016). WRKY transcription factors: Molecular regulation and stress responses in plants. Front Plant Sci 7, 2645. doi: 10.3389/fpls.2016.00760

Piasecka, A, Jedrzejczak-Rey, N, and Bednarek, P (2015). Secondary metabolites in plant innate immunity: conserved function of divergent chemicals. New Phytol 206, 948–964. doi: 10.1111/nph.13325

Pollastri, S, and Tattini, M (2011). Flavonols: old compounds for old roles. *Ann. Bot.* 108, 1225–1233. doi: 10.1093/aob/mcr234

Poupard, P, Parisi, L, Campion, C, Ziadi, S, and Simoneau, P (2003). A wound- and ethephon-inducible PR-10 gene subclass from apple is differentially expressed during infection with a compatible and an incompatible race of *Venturia inaequalis*. Physiol Mol Plant Pathol 62, 3–12. doi: 10.1016/S0885-5765(03)00008-0

Puhl, I, and Treutter, D (2008). Ontogenetic variation of catechin biosynthesis as basis for infection and quiescence of *Botrytis cinerea* in developing strawberry fruits. J Plant Dis Protect 115, 247–251. doi: 10.1007/BF03356272

Pühringer, H, Moll, D, Hoffmann-Sommergruber, K, Watillon, B, Katinger, H, and da Câmara Machado, M L (2000). The promoter of an apple Ypr10 gene, encoding the major allergen Mal d 1, is stress- and pathogen-inducible. Plant Sci 152, 35–50. doi: 10.1016/S0168-9452(99)00222-8

R Core Team. R: a language and environment for statistical computing. http://www.R-project.org/

Ranocha, P, Denancé, N, Vanholme, R, Freydier, A, Martinez, Y, Hoffmann, L, et al. (2010). Walls are thin 1 (WAT1), an Arabidopsis homolog of *Medicago truncatula* NODULIN21, is a tonoplast-localized protein required for secondary wall formation in fibers. Plant J. 63, 469–483. doi: 10.1111/j.1365-313X.2010.04256.x

Ren, X-Y, Vorst, O, Fiers, M W E J, Stiekema, W J, and Nap, J-P (2006). In plants, highly expressed genes are the least compact. Trends Genet 22, 528–532. doi: 10.1016/j.tig.2006.08.008

Rietz, S, Stamm, A, Malonek, S, Wagner, S, Becker, D, Medina-Escobar, N, et al. (2011). Different roles of Enhanced Disease Susceptibility1 (EDS1) bound to and dissociated from Phytoalexin Deficient4 (PAD4) in Arabidopsis immunity. New Phytol 191, 107–119. doi: 10.1111/j.1469-8137.2011.03675.x

Robinson, M D, McCarthy, D J, and Smyth, G K (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140. doi: 10.1093/bioinformatics/btp616

Rozen, S, and Skaletsky, H (2000). Primer3 on the WWW for General Users and for Biologist Programmers. In: Misener S., Krawetz S.A. (eds) Bioinformatics Methods and Protocols. Methods in Molecular Biology™, vol 132. Humana Press, Totowa, NJ. doi: 10.1385/1-59259-192-2:365

Schindelman, G, Morikami, A, Jung, J, Baskin, T I, Carpita, N C, Derbyshire, P, et al. (2001). COBRA encodes a putative GPI-anchored protein, which is polarly

localized and necessary for oriented cell expansion in Arabidopsis. Genes Dev 15, 1115–1127. doi: 10.1101/gad.879101

Slusarenko, A J, Fraser, R S S, and Loon, L C (2000). Mechanisms of Resistance to Plant Diseases. Dordrecht: Springer Netherlands.

Terefe-Ayana, D, Yasmin, A, Le, T L, Kaufmann, H, Biber, A, Kühr, A, et al. (2011). Mining disease-resistance genes in roses: functional and molecular characterization of the *Rdr1* locus. Front Plant Sci 2, 35. doi: 10.3389/fpls.2011.00035

Thimm, O, Blasing, O, Gibon, Y, Nagel, A, Meyer, S, Kruger, P, et al. (2004). MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J 37, 914–939. doi: 10.1111/j.1365-313X.2004.02016.x

Toljamo, A, Blande, D, Kärenlampi, S, and Kokko, H (2016). Reprogramming of Strawberry (*Fragaria vesca*) Root transcriptome in response to *Phytophthora cactorum*. PloS One 11, e0161078. doi: 10.1371/journal.pone.0161078

Tripathy B C, Pattanayak G K, ed (2012). Photosynthesis. Dordrecht: Springer Netherlands.

Tronchet, M, Balague, C, Kroj, T, Jouanin, L, and Roby, D (2010). Cinnamyl alcohol dehydrogenases-C and D, key enzymes in lignin biosynthesis, play an essential role in disease resistance in Arabidopsis. Mol Plant Pathol 11, 83–92. doi: 10.1111/j.1364-3703.2009.00578.X

van Loon, L C, Rep, M, and Pieterse, C M J (2006). Significance of inducible defense-related proteins in infected plants. Annu Rev Phytopathol 44, 135-162. doi: 10.1146/annurev.phyto.44.070505.143425

van Sandt, V S, Suslov, D, Verbelen, J-P, and Vissenberg, K (2007). Xyloglucan endotransglucosylase activity loosens a plant cell wall. Ann Bot 100, 1467–1473. doi: 10.1093/aob/mcm248

Vanholme, R, Demedts, B, Morreel, K, Ralph, J, and Boerjan, W (2010). Lignin biosynthesis and structure. Plant Physiol 153, 895–905. doi: 10.1104/pp.110.155119

von Malek, B, and Debener, T (1998). Genetic analysis of resistance to blackspot (*Diplocarpon rosae*) in tetraploid roses. Theor Appl Genet 96, 228–231. doi: 10.1007/s001220050731

Weiß, S, Bartsch, M, and Winkelmann, T (2017). Transcriptomic analysis of molecular responses in *Malus domestica* 'M26' roots affected by apple replant disease. Plant Mol Biol 94, 303–318. doi: 10.1007/s11103-017-0608-6

Whitaker, V M, Bradeen, J M, Debener, T, Biber, A, and Hokanson, S C (2010a). *Rdr3*, a novel locus conferring black spot disease resistance in tetraploid rose: genetic analysis, LRR profiling, and SCAR marker development. Theor Appl Genet 120, 573–585. doi: 10.1007/s00122-009-1177-0

Whitaker, V M, Debener, T, Roberts, A V, and Hokanson, S C (2010b). A standard set of host differentials and unified nomenclature for an international collection of *Diplocarpon rosae* races. Plant Pathol 59, 745–752. doi: 10.1111/j.1365-3059.2010.02281.x

Xu, Q, Wen, X P, and Deng, X X (2005). Isolation of TIR and non TIR NBS-LRR resistance gene analogues and identification of molecular markers linked to a powdery mildew resistance locus in chestnut rose (*Rosa roxburghii* Tratt). Theor Appl Genet 111, 819–830. doi: 10.1007/s00122-005-0002-7

Yamamoto, M, Nakatsuka, S, Otani, H, Kohmoto, K, and Nishimura, S (2000). (+)-catechin acts as an infection-inhibiting factor in strawberry leaf. Phytopathology 90, 595–600. doi: 10.1094/PHYTO.2000.90.6.595

Yang, Y-X, Ahammed, G J, Wu, C, Fan, S-Y, and Zhou, Y-H (2015). Crosstalk among Jasmonate, Salicylate and Ethylene signaling pathways in plant disease and immune responses. Curr Protein *Pept Sci.*16, 450–461. doi: 10.2174/1389203716666150330141638

Zhang, S-H, Yang, Q, and Ma, R-C (2007). *Erwinia carotovora* ssp. *Carotovora* infection induced "defense lignin" accumulation and lignin biosynthetic gene expression in chinese cabbage (*Brassica rapa* L. ssp *pekinensis*). J Integr Plant Biol 49, 993–1002. doi: 10.1111/j.1672-9072.2007.00478.x

Zhou, N, Tootle, T L, Tsui, F, Klessig, D F, and Glazebrook, J (1998). PAD4 functions upstream from salicylic acid to control defense responses in Arabidopsis. Plant Cell 10, 1021–1030. doi: 10.1105/tpc.10.6.1021

6. Additional results

Besides the data presented in the previous chapters, some additional results were generated during this thesis, which were not fully analysed but should be reported here as valuable resources for further studies.

6.1. The draft genome of the rose genotype 88/124-46

Even before the thesis started the publication of a reference sequence for the rose genome was announced, which was, however, only published very recently (Hibrand *et al.*, 2018). Due to this fact, we started working on the annotation of a draft genome sequence of the diploid genotype 88/124-46 (further referred to as 88 genome), which was chosen because it contains the Rdr1-locus (Drewes-Alvarez, 1992; Malek and Debener, 1998). The assembly of the sequence itself was mainly performed by Jonathan Featherston (Agricultural Research Council). Due to the fact that we received early access to the reference genome of the rose variety ‘Old Blush’ (Hibrand *et al.*, 2018), the work on the 88 genome was not followed up. Nevertheless, the sequence can be used for genotypic specific analysis, primer development or for comparative analysis.

Material and methods

The genome sequence

For the sequencing, DNA was extracted of young leave material according to (Kobayashi *et al.*, 1998). Sequencing of a short insert library as well as a mate paired library was performed by our cooperation partners of the Biotechnology Platform from the Agricultural Research Council (Pretoria, South Africa) using an Illumina HiScan SQ and an Illumina HiSeq 2500. These data were completed by one run of Single molecule real time (SMRT) sequencing performed by GATC (Karlsruhe, Germany) on a PacBio RS (Pacific Biosciences, Menlo Park, USA) with an insert size of 5-10kb.
Genome assembly of these reads was performed by Jonathan Featherston and Jasper Rees (Agricultural Research Council, Pretoria, South Africa).

Annotation

Gene prediction was performed similar to the procedure used for genome sequence of *D. rosae* (Chapter 2). The MAKER 2.3.1.8 pipeline (Cantarel *et al.*, 2008) was used with three different *de novo* gene prediction programs: GeneMark-ES (Lomsadze *et al.*, 2005), Augustus (Stanke *et al.*, 2006) and SNAP3 (Korf, 2004). All prediction programs were trained with the analysed assembly of the genotype 88/124-46. GeneMark-ES is a self-training prediction program. Augustus was automatically trained during analysis with BUSCO 1.2 (Conesa *et al.*, 2005) on the bases of the identified orthologs. SNAP3 was trained by three rounds of hint-based MAKER prediction without any other prediction tool. The following sources were used as evidence for the MAKER annotation: a collection of full length ESTs from different rose genotypes (Koning-Boucoiran *et al.*, 2015), an assembly of all MACE and RNAseq data presented at this thesis generated using *de novo* assembly algorithm of CLC Genomics Workbench 9.0.1 (Qiagen, Hilden, Germany), the public available gene models and predicted protein sequences of the *Fragaria vesca* genome v.2.1 (Tennessen *et al.*, 2014) and the predicted protein sequences of *Prunus persica* (V. 2.0) (Verde *et al.*, 2017) and *Malus domestica* (Velasco *et al.*, 2010). As an additional tool, tRNAscan (Lowe and Eddy, 1997) was used for the prediction of tRNA-genes.

Functional annotation was done by means of Blast2Go 3.3 (Conesa *et al.*, 2005) using BLASTx with a subset of all plant sequences of the NCBI non-redundant (NR) protein database (E-value cutoff e-10) including a Gene Ontology (GO) term annotation (Ashburner *et al.*, 2000).

Additional analysis

As benchmark for the completeness of the gene space included in the genome sequence and the annotation, the BUSCO 1.2 pipeline (Simao *et al.*, 2015) was used.

Results

Due to the fact that generating a rose reference sequence took longer than expected, the 88 genome was used for annotation.Table 6.1 contains the key data describing the assembly. The assembly comprises 27,094 scaffolds with a total length of

approximately 237 Mbp and a N50 value of 20,325. The majority of these scaffolds are larger than 1000 bp with the largest scaffold having a length of 243.9 kbp and 11.47% of the assembly comprising ambiguous bases which are mainly included as gaps during the procedure of scaffolding.

Table 6.1: Key data of the assembly of the genotype 88/124-46

Parameter	Value
Total number of scaffolds	27,094
Total length of all scaffolds	236,696,190 bp
Number of scaffolds larger 1000 bp	25,698
Total length of scaffolds larger 1000 bp	235,358,791 bp
Largest scaffold	243,924 bp
N50	20,325 bp
Fraction of ambiguous bases	11.47%

For the annotation, the MAKER pipeline (Cantarel *et al.*, 2008) was used with additional supporting evidences of related species and an assembly of the presented transcriptomic data. The number of predicted genes differed noticeable depending on the prediction tool (Table 6.2). SNAP (Korf, 2004) and Augustus (Stanke *et al.*, 2006) predicted similar numbers of 34,523 or 38,008 genes, respectively. GeneMark (Lomsadze *et al.*, 2005), with 64,863, predicted a higher number of genes than the other two tools. By comparing the output of different prediction tools and including the supporting evidences MAKER generated 25,868 high quality annotations. In addition, 466 non-coding tRNA genes were predicted. The assembly of the 88 genome as well as the results of the MAKER pipeline are included in the electronical appendix (DVD 1) as electronic supplementary material ES1.

Table 6.2: Results of the MAKER annotation pipeline

Annotation tool	Number of gene models
GeneMark	64,863
SNAP	34,523
Augustus	38,008
High quality annotation	25,868

To estimate the proportion of the gene space which is included in the genome as well as in the high quality annotations generated with MAKER, the BUSCO pipeline was used. As can be seen in Table 6.3, approximately 80% of the BUSCO orthologs were identified as complete sequence in the assembly and in the annotations, and only 13.9% or 12.7%, respectively, were not detected. The rest, 6.4% and 6.9%, are only detected as fragmented sequences. The results of the BUSCO pipeline are included in the electronical appendix (DVD 2) as electronic supplementary material ES2.

Table 6.3: Results of the BUSCO pipeline for the 88 genome assembly and the high quality annotations generated with the MAKER pipeline.

	Genome [%]	MAKER high quality annotation [%]
Complete	79.9	80.4
Fragmented	6.4	6.9
Missing	13.9	12.7
Total BUSCO groups searched	1440	1440

6.2. Additional MACE and RNAseq data

In addition to the transcriptomic data presented in chapter 5, MACE-data of the incompatible interaction of *D. rosae* with the resistant genotype 91/100-5 and a transgenic PC genotype (muRdr1A) carrying the RGA1 gene of the Rdr1-locus (Menz *et al.*, 2017) were generated. Additionally, PC samples of leaves which were frozen directly after the sampling were sequenced to gain insight into general stress responses derived from the inoculation procedure.

To have the possibility to compare conventional RNAseq with the MACE approach, the same RNAs of PC respectively 91/100-5 inoculated with *D. rosae* (72hpi) used for the MACE approach were used for the application of conventional RNAseq.

Material and methods

Start samples

In parallel to all three inoculation experiments described in chapter 5, samples of uninoculated PC leaves were frozen in liquid nitrogen directly after the extraction from the plant (Start) (Table 6.4). These differ from the 0hpi control samples in a way that they only experienced the stress through the cut but not the inoculation procedure.

Inoculation of 91/100-5 and muRdr1A

In addition to the inoculation of the second and the third biological repetition of the inoculation experiments described in chapter 5, leaves of the resistant genotypes 91/100-5 and the transgenic genotype muRdr1A (sample name RGA1 generated only for BR2) were inoculated with the same spore solution as the PC leaves. Additionally, a fourth inoculation (BR4) experiment was performed solely with the genotype 91/100-5 in a similar way as the previously performed inoculations. Samples of these incompatible interaction systems were taken at 24hpi and 72hpi.

Sampling and RNA extraction as well as application of the MACE procedure were performed as described in the material and methods part of chapter 5.

RNA samples of the 72 hpi time point from the inoculation experiments of the compatible interaction of PC with *D. rosae* (BR1/BR2/BR3: PC+DR_72hpi), described in chapter 5, and the previously mentioned incompatible interaction of the genotype 91/100-5 (BR2/BR3/BR4: 91+DR_72hpi) were used for an RNAseq approach. For a better overview, Supplementary Table 1 contains the performed inoculation experiments and the data sets derived from these experiments. Library preparation and paired-end sequencing was performed by GATC Biotech (Konstanz, Germany).

Further bioinformatic processing of the additional MACE and RNAseq-data was performed as described in the material and methods part of chapter 5.

Results

MACE application of the additional samples resulted in comparable amounts of sequenced reads ranging from approximately 7.6 to 14.5 million reads per library (Table 6.4). Also, the mapping results are similar to the previously mentioned MACE-data and range from 87.44% to 92.59% of reads uniquely mapped to the genome sequence of the rose variety ‘Old Blush’ (Hibrand *et al.,* 2018). The mapping results, normalised expression values as well as the statistical data of relevant comparisons are included in the electronical appendix (DVD 2) as electronic supplementary material ES3.

The RNAseq approach resulted in approximately 60 to more than 100 million reads per library in a much higher read out-put than the MACE-approach (Table 6.5). 22.3 to 41.2 million reads mapped in pairs and the majority of them mapped uniquely in the rose reference sequence. The mapping results, normalised expression values as well as the statistical data of the comparisons are included in the electronical appendix (DVD 2) as electronic supplementary material ES3.

Table 6.4: Overview of the results of the additional MACE-data

Sample	Repetition	No. of Reads	Average length [bp]	Total reads mapped [No. | %]		Unique mapped reads [No. | %]	
Start	BR 1	10,772,127	89.2	9,873,919	| 91.66	9,687,773	| 89.93
	BR 2	14,519,586	89.4	13,408,923	| 92.37	13,179,153	| 90.79
	BR 3	13,528,517	110.6	12,726,539	| 94.07	12,515,677	| 92.51
91+DR 24 hpi	BR 2	11,012,063	88.0	9,917,069	| 90.06	9,645,232	| 87.59
	BR 3	12,088,992	109.5	10,866,714	| 89.89	10,570,525	| 87.44
	BR 4	7,622,351	111.0	6,886,414	| 90.35	6,713,987	| 88.08
91+DR 72 hpi	BR 2	8,291,465	88.0	7,554,600	| 91.11	7,364,416	| 88.82
	BR 3	11,439,079	110.8	10,282,087	| 89.89	10,008,852	| 87.50
	BR 4	11,374,829	110.3	10,186,486	| 89.55	9,940,741	| 87.34
RGA1+DR 24 hpi	BR 2	11,025,588	87.5	10,069,661	| 91.57	9,815,125	| 89.02
RGA1+DR 72 hpi	BR 2	7,752,983	87.2	7,148,605	| 92.20	6,964,662	| 89.83

Table 6.5: Overview of the results of the additional RNAseq-data

Sample	Repetition	No. of Reads	Average length [bp]	Reads mapped in pairs	Total pairs mapped [No. | %]	Unique pairs [No. | %]
PC+DR 72 hpi	BR 1	61,495,172	124,9	48,036,520	24,018,260 | 78,11	23,118,631 | 75.19
	BR 2	60,004,702	124,8	44,513,724	22,256,862 | 74,18	21,227,003 | 70.75
	BR 3	96,457,220	124,9	77,749,394	38,874,697 | 80.61	38,136,422 | 79.07
91+DR 72 hpi	BR 2	103,907,224	124,3	82,433,066	41,216,533 | 79,33	40,490,037 | 77.93
	BR 3	90,462,452	124,6	72,850,852	36,425,426 | 80,53	36,015,205 | 79.62
	BR 4	99,692,850	124,6	79,645,866	39,822,933 | 79,89	39,242,212 | 78.73

6.3. Extraction of sequences from *Podosphaera pannosa*

P. pannosa is the causing agent of the powdery mildew disease, one of the most severe diseases on roses, especially for the greenhouse production of cut roses (Debener and Byrne, 2014). Surprisingly, except for some internal transcribed spacer (ITS) and ribosomal sequences, no sequence information is available for this fungus in the NCBI-NR database yet. Due to this fact, the MACE-data generated for the compatible interaction of the rose genotype PC with the *P. pannosa* were used to isolate sequences of the fungus which could be a starting point for further analysis of the fungus.

Material and methods

MACE-data of the three biological repetitions of the compatible interaction of PC with a greenhouse mixture of *P. pannosa* (Chapter 5) were the source for extracting sequence information of the fungus. Figure 6.1 illustrates the procedure used for the extraction of these sequences.

To identify the majority of reads belonging to the plant, all reads of the three time points were mapped to the 88 genome using the CLC genomics workbench 9.0 (Qiagen, Hilden, Germany) with the following parameters: Mismatch cost=2, insertion cost=3, deletion cost=3, length fraction=0.9, similarity fraction=0.9. Unmapped reads were assembled with the *de novo* assembly algorithm implemented in the CLC genomics workbench (Qiagen, Hilden, Germany), using a word size of 23 and a bubble size of 50. These values were automatically chosen by the assembly algorithm. Contigs were automatically updated after mapping the reads back to the generated contigs with the above mentioned mapping parameters except for the length fraction and similarity fraction which were both set to 0.95. The resulting contigs were aligned to the genome sequence and the predicted transcriptome of the 88 genome by means of the BLASTn algorithm (E-Value cut-off 1E-10) (Camacho *et al.*, 2009). Matching contigs were discarded. The remaining contigs were aligned to subsets of the NCBI NR database containing solely fungal or plant sequences respectively with the BLASTx algorithm (E-Value cut-off 1E-10). Sequences matching only the plant database were discarded and those matching only the fungi database were assumed to be of fungal origin. Sequences matching both databases were used

for BLASTx alignment with the complete NR database. Contigs with the best match coming from fungal sequence were also assumed to be of fungal origin, which implies that they belong to *P. pannosa*.

Functional annotation was done by means of Blast2Go 3.3 (Conesa *et al.*, 2005) using BlastX against a subset of all fungal sequences of the NCBI NR protein database (E-value cutoff e-10) including a Gene Ontology (GO) term annotation (Ashburner *et al.*, 2000)

To identify full-length transcripts among the identified sequences of *P. pannosa* the web version of the Full_lengtherNEXT package (Lara *et al.*, 2007) was used.

Additionally, the sequences were used for a tBLASTx and a BLASTn (E-value cutoff e-10) alignment to the predicted mRNA sequences of the related fungus *P. xantii* (Vela-Corcía *et al.*, 2016).

Results

Due to the fact that no genome sequence or any other sequence information of *P. pannosa* are available, the procedure illustrated in Figure 6.1 was used to identify sequences of fungal origin amongst the MACE-data. This approach resulted in 5056 contigs potentially originating from *P. pannosa*. In addition to the BLAST-derived description, Blast2Go assigned GO-terms to 4425 contigs and E.C. numbers to 1188 of them. As expected, due to the fact that the sequences are assembled from MACE-tags representing only the 3’ end of a molecule, only 213 of the contigs were identified as full-length sequences by the Full_lengtherNEXT package. The web server also identified 333 sequences, which might be chimaeras. A BLAST alignment of the potential *P. pannosa* sequences with the related cucumber powdery mildew fungus *P. xantii* resulted in 4374 matches for the tBLASTx algorithm and 3892 matches for the BLASTn algorithm. The assembled contigs and annotations as well as the results of Full_lengtherNEXT and the BLAST are included in the electronical appendix (DVD 2) as electronic supplementary material ES4.

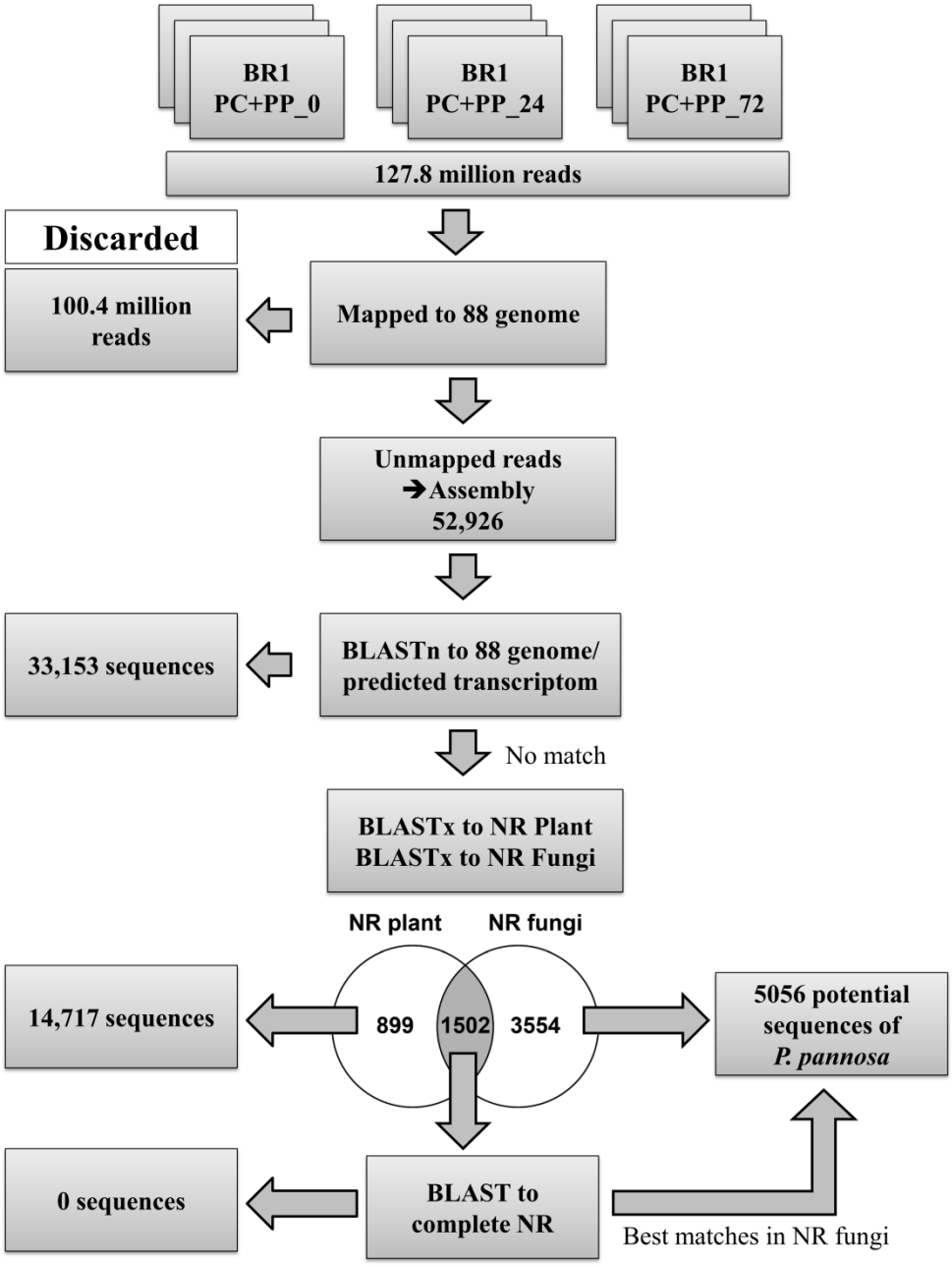


Figure 6.1: Overview of procedure applied for the isolation of sequences belonging to *P. pannosa*. The MACE data of the compatible interaction between PC with *P. pannosa* acted as the source for the extraction of the fungal sequences. These reads were at first mapped against the 88 genome and the unmapped reads were used for a *de novo* assembly. The assembled contigs were aligned to the 88 genome and its predicted transcriptome via the BLASTn algorithm (E-value threshold 1E-10). Matching sequences were discarded. The remaining contigs were aligned with subsets of the NCBI NR-base for both plant and fungi by means of the BLASTx algorithm (E-value threshold 1E-10). Sequences solely matching the NR plant database were discarded, those matching solely the NR fungi database were assumed to originate from *P. pannosa*. The remaining contigs matching both databases were used for another alignment via BLASTx to the complete NR-base. Sequences resulting in a best match with fungal sequences were further assumed to belong to *P. pannosa*.

6.4. Establishment of a P-starvation experiment

Phosphorus (P) is an essential macronutrient and fulfils various physiological functions in plants. It is part of different macromolecules and plays an important role in the energy transfer and sugar or starch translocation. It is also involved in the regulation of many metabolic processes and is a key factor in signal transduction. The P supply of a plant also influences its defence reaction, but the effect of P deficiency on this mechanism is inconsistent across pathosystems. The results of several experiments suggest that a P application can either decrease or increase the susceptibility of a plant (Prabhu *et al.*, 2007; Hubert *et al.*, 2012). At the beginning of this project, the main focus was on the influence of P-starvation on the defence response of roses. To investigate this question, a hydroponic system was established and based on typical starvation symptoms like reduced leaf expansion as well as the root shoot/root ratio, treatments of high and low P-supply were defined. Especially the low P-supply treatment is critical because the reduced leaf area as result of the P-starvation can make following inoculation experiments impossible.

Material and methods

Plant Material

For the establishment of P-starvation experiments, 12 genotypes of the diploid 97/7 population (Linde *et al.*, 2006) were chosen. Six of these genotypes (97/7-81, -123, -131, -189, -230, -244) were identified as double resistant to powdery mildew and black spot in previous tests, the other six genotypes (97/7-39, -69, -126, -183, -214, -239) were characterised as double susceptible to both pathogens (Linde *et al.*, 2006, unpublished data).

The source material of all experiments are plants of the genotypes grown in pots under greenhouse conditions in soil “Typ 0 Nullerde” (Einheitserdewerke Patzer, Sinntal-Altengronau, Germany), without any additional fertiliser added to the substrate. After the plants showed visible P-starvation symptoms, they were fertilised once a week with a customised fertiliser (Supplementary Table 2) containing 10 µM KH_2PO_2 as P source.

Determination of the P-treatment

Cuttings of the above described plants were rooted in nutrient solution (Supplementary Table 3) without any KH_2PO_2 added for four weeks under a translucent tent and nearly 100% humidity. The nutrient solution was renewed once a week. To prevent mold infections, the plants were treated with a contact fungicide once directly after transfer to the nutrient solution and a second time after two weeks of rooting. For acclimatisation, successfully rooted cuttings were transferred into 10 l buckets containing nutrient solution without KH_2PO_2 added and under continuous air supply to the solution. Genotypes were randomly arranged in the bucket. After one week, plants were transferred into the nutrient solution containing different amounts of KH_2PO_2 as P-source (0, 1, 3, 5, 10, 100 µM). Plants were grown in this solution for 8 weeks under greenhouse conditions with renewing the solution once a week. After this period, the leaf length and width of five leaves of every genotype cultured under the different P-amounts was measured (Supplementary Figure 1). Additionally, the total root and shoot mass was separated and dried in open plastic bags at 38°C for three weeks to determine the root-shoot ratio based on the dry masses (Supplementary Figure 2). The procedure was performed in three biological repetitions in two experiments. A 10 µM treatment was included only in the second experiment, thus only data for two biological receptions were collected.

Results

A hydroponic system was established to investigate the influence of P-starvation on the defence response. To determine a P-treatment where clear starvation symptoms are visible but inoculation experiments are still feasible, six different P-concentrations in the nutrient solutions were tested. Figure 6.2 and Figure 6.3 depict the results of the three starvation indicators leaf length, leaf width and root/shoot ratio as a box plot (Electronic supplementary material ES5). The leaf area increases constantly with higher P-supply with the highest leaf length and width by treatment with 100 µM KH_2PO_2 as P-source in the nutrient solution. The same effect is also visible for the root/shoot ratio. The difference between 100 µM KH_2PO_2 and the other P-treatments is even more visible here. For further starvation experiments the recommendation is to use 3 µM as low P-treatment and 100 µM as high P-treatment because a clear

difference in the starvation symptoms is visible between the two treatments. Starvation symptoms by treatment with less than 3 µM KH_2PO_2 in the nutrient solution were too severe on some genotypes. They produced no or not enough leaf area to perform inoculation experiments.

Another result of this experiment is that the time period of the experiment can be reduced to prevent contaminations with mold infections or powdery mildew. Already after three to five weeks of culture in the nutrient solution, a clearly reduced growth under the 3 µM KH_2PO_2 was visible (Supplementary Figure 3).

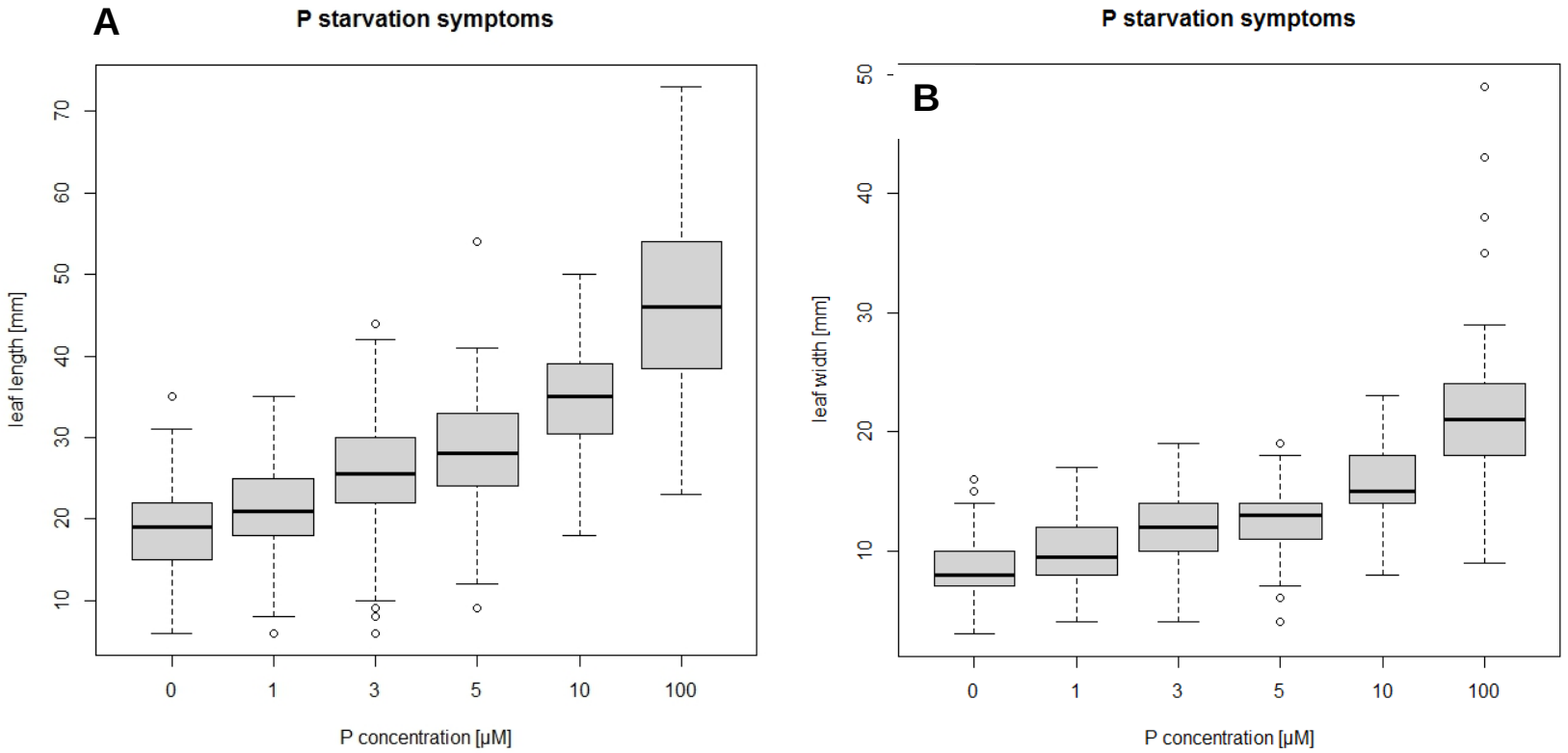


Figure 6.2: Boxplot of growth parameter of leaves as indicators for P-starvation. The leaf length (A) and width (B) under different P-concentrations in the nutrient solution are depict. The boxplot contains data of 12 different genotypes per treatment and three biological repetitions.

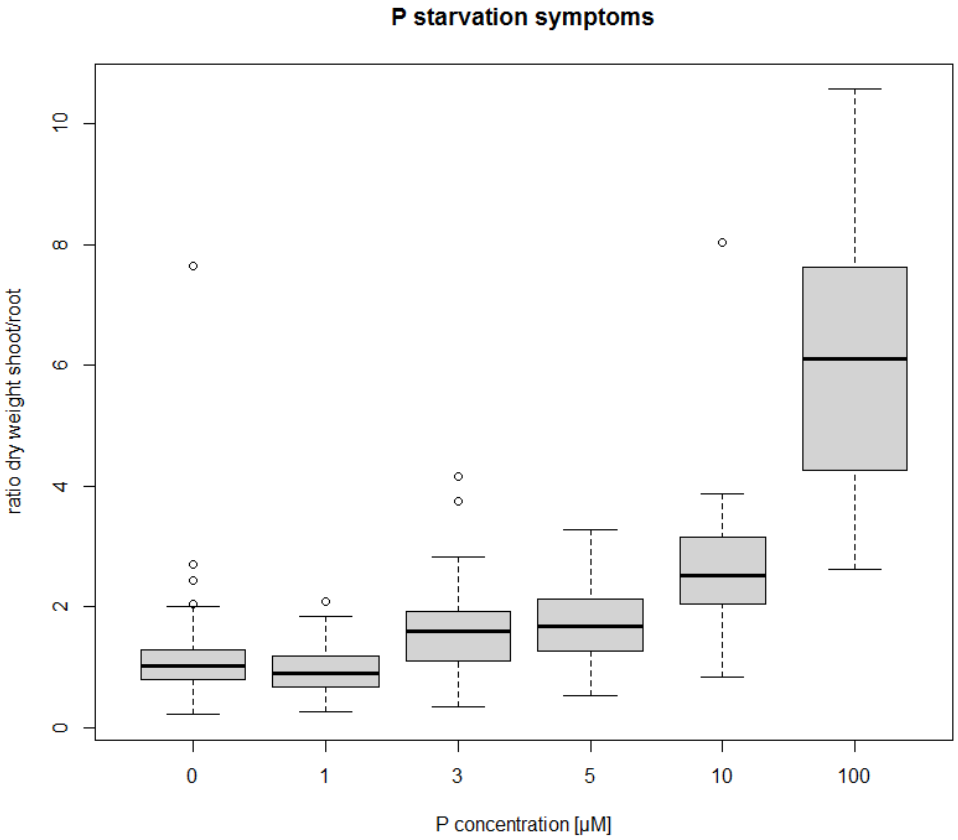


Figure 6.3: Boxplot of the shoot/root-ratio as indicator P-starvation. The ratio of the dry weight of the shoot and root under different P-concentrations in the nutrient solution is depicted. The boxplot contains data of 12 different genotypes per treatment and three biological repetitions.

7. General discussion

The main goal of this study was to generate transcriptomic data to analyse the response of the rose leaf transcriptome to biotic stress. Due to the complexity of the highly heterozygous tetraploid genome of the rose and the mixture of plant and fungal sequences in the transcriptomic data, the *de novo* analysis was much more complicated than expected. Only utilising the generated genome sequence of *D. rosae* and the reference sequence of the rose genome made it possible to characterise and compare the transcriptomic responses to *D. rosae* and *P. pannosa* and to finally point out important defence related mechanisms. The main results and the conclusions derived from the analyses of these resources have already been described and discussed in the previous chapters 3 to 6. The following chapter will focus on the relationship between these findings and how to use the generated results in further experiments.

7.1. The draft genome sequence of *D. rosae*, a typical untypical fungal genome

Even if black spot is the most important disease in growing garden and landscape roses and resistance to this pathogen is one of the main goals in rose breeding (Debener and Byrne, 2014), so far no genomic sequence for *D. rosae* was available.

The presented draft genome sequence of the black spot fungus is still highly fragmented with 2457 scaffolds. One of the major problems in assembling this genome is the untypically large proportion of sequence duplication found during the analysis of the genome sequence (chapter 2). Including data of longer reads, e.g. those generated with Oxford Nanopore or PacBio sequencing (Li *et al.*, 2017a), could further improve the genome sequence through closings of gaps and assembling of previously unassembled contigs if they bridge highly similar regions of the genome.

Together with transposable elements, the genome duplication is one of the major reasons for the enlargement of the genome leading to a predicted genome size of 72.5 and 91.4 Mb. So far our efforts to estimate the genome size experimentally by flow cytometry were not successful. However, the application of normal staining solution like 4′,6-Diamidin-2-phenylindol or propidium iodide of staining nuclei within

spores is possible (Supplementary Figure 4) so that other methods like image cytometry (Kullman and Teterin, 2006) can be used to validate the estimation based on the sequence information.

The genome duplication is an interesting phenomenon which could be one of the driving forces in the evolution of *D. rosae* like it was shown in yeast, *Microsporidia* or *Rhizopus oryzae* (Ma *et al.*, 2009; Albertin & Marullo, 2012; Williams *et al.*, 2016). Surprisingly, the genome of the closely related species *M. Brunnea* showed no signs of larger duplicated regions according to the BUSCO analysis (Chapter 3). It would be very interesting to analyse the genomes of other *Diplocarpon* species like *D. Mali* or *D. earlianum* which use other members of the Rosaceae family as host.

Furthermore, different isolates of *D. rosae* showed differences in the analysed pairs of paralogous sequences. This could be an indication for differences in the proportion genome that is duplicated. For further analyses, the isolates I001 and R6 might be of particular interest, since they showed the strongest difference in the duplication pattern. R6 was included in a previous study of AFLP markers and rDNA sequencing to analyse the diversity of different *D. rosae* isolates (Blechert, 2005). In this study the isolate was called *D. rosae* ‘101’. Surprisingly, R6 did not clustered together with other isolates of *D. rosae*, it was as related to the other *D. rosae* isolates as other analysed *Diplocarpon* species like *D. mali, D. earlianum* or *D. espili*. It might be that the differences in the duplicated proportion of the genome are one of the reasons that differentiate R6 from the other isolates.

The genome sequence presented here is generated from DNA isolated from *in vitro* conidia and mycelium. According to our observations, conidia derived from *in vitro* cultures are unable to infect leaves (unpublished data). One of the hypotheses explaining this phenomenon could be that supernumerary chromosomes are lost during the *in vitro* culture, comparable to the loss of plasmids in bacterial cultures. As was shown, these supernumerary chromosomes can be crucial for the virulence of fungi if they contain important virulence factors or effector genes (Yoshida *et al.*, 2009; Ma *et al.*, 2010). Resequencing of DNA generated from spores washed from infected leaves could clarify whether larger parts of the genome are missing in the presented assembly.

Nevertheless, the draft genome is still a good working tool for further analysis. The BUSCO analysis indicates that the sequence contains the complete gene space and as chapter 3, shows it enables us to analyse genes like effectors or vir factors. The large number of SSR identified in the genome can be used for large- scale marker application needed for diversity analysis and to investigate the phylogenetic position of *D. rosae*. Moreover, the genome sequence enables us to identify the mating type locus, which was only partially isolated yet (Lühmann, 2010), to investigate the sexual reproduction of this fungus.

7.2. The predicated secretome and effector content of *D. rosae*

Besides the necessity to generate sequence information of *D. rosae* to separate fungal and plant reads in the performed transcriptomic analysis, one of the main reasons for sequencing the genome of *D. rosae* was to identify potential virulence factors and effectors in the secretome of the pathogen.

To increase prediction accuracy, a combination of different bioinformatic approaches was applied to predict the secretome of the fungus. So far there is no consensus in the scientific community about the best combination of tools for this task. The approach utilised in the presented study is derived from the pipeline used by the Fungal Secretome Knowledgebase (FunSecKB) (Meinken *et al.*, 2014) which collects data about secretome of all sequenced fungi except for the tools TargetP and WolfPSort. The results of these programmes were only reported as additional information and not as disqualifiers because TargetP uses the same algorithm as the included tool SignalP and WolfPSort was reported to have a low sensitivity (Sperschneider *et al.*, 2015b). However, even by applying stringent criteria and combining different prediction tools false positives as well as false negatives are unavoidable in *in silico* analysis. Due to this fact, for the analysis of individual factors of the secretome experimental approaches like GFP-fusion or secretion trap assay (Lee and Rose, 2012) are needed to validate the subcellular localisation of a secreted protein. In addition, proteins translocated by the unconventional secretion pathway are unpredictable with the established methods, thus these proteins are not included in the presented secretome of *D. rosae*. A way of identifying additional secreted proteins is a proteomic analysis of the culture medium the fungus is growing

in like it was applied for *Fusarium graminearum* (Yang *et al.*, 2012), even if this system is quite artificial.

Furthermore, interesting factors which were only found in the analysis of the whole genome (chapter 2) but are not included in the predicted secretome (chapter 3), like for instance the LysM domain containing proteins, can still be targets for further studies and should not be excluded solely on the fact that no secretion was predicted for these proteins.

The analysis of the secretome revealed a large number of plant CWDEs which might include important vir factors. Especially the strongly expressed genes coding for pectin-degrading enzymes are interesting candidates. It is quite surprising that only very few expressed lignin-degrading enzymes were identified taking into account that the transcriptomic analysis of the plant defence indicated that an up-regulation of the lignin pathway is one of the specific responses to *D. rosae* (chapter 4). This might indicate that the up-regulation of this pathway has no effect in restricting the fungal growth or its nutrient supply so that the fungus needs no counter-measures. Detailed analysis of the response of the lignin pathway and the lignin degrading genes of the fungus e.g. in an incompatible interaction where the growth of the fungus is restricted can give insights into the importance of this phenomenon.

For the pathogen-host interaction, the most important class of secreted proteins are effector proteins because of their dual function as important vir factors and avr factors (Bent and Mackey, 2007). The effector of *D. rosae* which is recognised by the *RGA1* gene of the Rdr1 locus is of particular interest. It must have a crucial role for the fungus because the Rdr1 locus mediates a broad spectrum resistance to many different isolates of several pathogenic races (Menz *et al.*, 2017) and not only to one race as it is expected regarding the gene-for-gene hypothesis (Flor, 1971).

In the genome of *D. rosae*, 251 effector candidates were predicted and 52 of them matched all the criteria of small secreted proteins, contain a Y/F/WxC motif (Godfrey *et al.*, 2010) and are predicted with the tool EffectorP (Sperschneider *et al.*, 2016). The fact that these candidates are predicted by different approaches increases the reliability of the prediction, making this set of candidates the best starting point for further analysis. These candidates are of particular interest because of the occurrence of the Y/F/WxC motif which was so far only reported for predicted

effectors of obligate biotrophic fungi (Godfrey *et al.*, 2010; Duplessis *et al.*, 2011; Saunders *et al.*, 2012). So far this is the first time that this class of effectors was found in a hemibiotrophic pathogen, indicating that this motif might be a useful tool to identify new effectors of pathogens with different lifestyles and not just biotrophic fungi. Analysing this motif in known effectors of other fungi would be an easy approach to get additional hints whether this motif is really directly associated with effector proteins or whether its occurrence is only a coincidence derived from the fact that effectors are often cysteine-rich and the comparable simple structure of motif. If the motif is really associated with effector proteins, it could improve the effector prediction in fungi like the RxLR-motif in oomycetes (Anderson *et al.*, 2015).

However, the set of effector candidates presented here should be analysed in detail to identify the functional avr gene corresponding to the *RGA1* gene of the Rdr1 locus, to gain a deeper understanding of the pathogenesis of the fungus. Transient expression of the candidates in resistant plants containing the Rdr1-locus would be an uncomplicated approach to identify the functional avr gene of the isolate DortE4. The expression of the effector should trigger an ETI reaction leading to necroses due to the HR. A similar approach would be coinfiltrating *Nicotiana benthamiana* leaves with the effector candidates and the resistance gene to induce a HR. This system was already successfully applied for the interaction of *Avr2* of *F. oxysporum* and corresponding R-gene *I-2* of tomatoes (Ma *et al.*, 2012).

Another crucial step in analysing the function of the effector candidates is the generation of knock-out or gain-of-function mutations of the fungus. So far no transformation system for *D. rosae* is established, but *Agrobacterium tumefaciens* mediated transformation is a common approach in transforming fungi and was already applied in the related species *M. brunnea* (Jiang *et al.*, 2014; Li *et al.*, 2017b). Different outcomes of these experiments are possible. The knock-out of an effector with avirulence function should prevent its detection through the R-protein and thus the isolate should become virulent. Additionally, the transformation of a virulent isolate with an avr factor should lead to detection through the R-protein and a resistance reaction. A knock-out of effectors with virulence function or other virulence factors should at least reduce the virulence of the pathogen.

Another interesting aspect includes comparing the effector content of different isolates of *D. rosae,* e.g. by resequencing their genomes. As mentioned above, the

isolate R6 could be a good starting point. This isolate can break the Rdr-1 mediated resistance indicating that it either has an additional effector which efficiently interferes with the plant defence or the effector detected by the R-protein is missing in the genome of this isolate. In both cases data of the effector content can give important information about the co-evolution of the pathogen with its host.

Previous unpublished experiments indicate that different rose genotypes and populations of our collection contain additional resistance genes besides the Rdr1-locus including race-specific resistances. These resources can be used to identify different avr factors with the above mentioned approaches. Comparing these different effector genes might help us to understand why some effectors, like the one recognised by the Rdr1-locus, mediate a broad spectrum resistance whereas other effectors are responsible for a race specific resistance.

Besides the analysis of pathogenesis-related processes of *D. rosae,* the knowledge about vir and avr factors of the fungus can also be used to improve resistance breeding of roses. Resistance genes recognising avr factors shared by isolates of several races are more durable than those resistance genes which recognise race specific effectors. With the knowledge of the effector content of *D. rosae*, such durable resistance genes can be better identified. In addition, effectors with vir function can be used to identify their targets in the plant cell by methods like yeast-two-hydride or affinity chromatography. The identification of varieties of these plant proteins unaffected by the vir factor can improve the tolerance against the pathogen.

All in all, the presented set of potential virulence factors and effector candidates represent the starting point for various analyses which can lead to a deeper understanding of the pathogenesis of *D. rosae* and its interaction with its host plant. Furthermore, these findings can help improving resistance breeding against this pathogen.

7.3. Application of the MACE-technique in roses

At the beginning of this study, no sequence information for the rose or the pathogens was available. Due to this fact, a *de novo* approach was used to analyse the transcriptomic data generated with the MACE approach. Due to the genomic

complexity of roses and the combination of plant and fungal sequences in the samples, the *de novo* analysis resulted in many differentially expressed sequences which could not be verified in further expression analyses. Different reference sequences were assembled from the reads derived from the MACE-approach and the RNAseq data, but all of them had problems with false positive results. There were two main reasons for these problems. Due to the mixture of plant and fungal sequences in the data sets, many contigs in the assembly were of fungal origin, which showed expression only in inoculated samples but no expression in the control treatment where no fungus was present. In addition, the assemblies were highly redundant due to the complexity of the rose genome. Both genotypes used for this study (PC and 91/100-5) are tetraploid and highly heterozygous, meaning that each gene can appear in up to four different alleles in each genotype. Additionally, many genes in the rose genome appear in large families, increasing the redundancy of very similar sequences. This redundancy leads to ambiguous mappings and genes which were erroneously detected to be only expressed in one of the two genotypes.

Another problem was the annotation of the assembled contigs. Due to the fact that the MACE approach only produces small tags of the 3' end, a transcript which often contains UTR regions, many sequences could not be annotated because the included coding region was not large enough to find homologous sequences in the databases. Only when the RNAseq data were included into the generation of the assembly, enough meaningful annotations were assigned to the contigs.

The described problems indicate that a *de novo* approach without a reference sequence is not applicable in such a complex system like roses and that good reference sequences are essential to generate meaningful results. Due to this fact, a satisfying analysis of MACE-data was only possible by utilising the genome of *D. rosae* presented in this thesis and the newly available rose reference sequence.

7.4. Generation and annotation of a rose genome sequence

As described above, a genomic sequence of the rose genome was needed for the analysis of the generated transcriptomic data. Due to this fact, the genome of the rose genotype 88/124-46 was sequenced and annotated (chapter 6.1). The genotype was chosen because of its diploid nature which reduces the genomic complexity

drastically and the presence of the Rdr1 resistance locus in its genome. With more than 27,000 scaffolds and 11.47% of ambiguous bases, the generated genome is still very fragmented (Table 6.1). Nevertheless, the predicted transcriptome contains more than 80% of the highly conserved orthologs used as a benchmark in the BUSCO pipeline (Table 6.3).

In parallel to the work on the 88 genome, we gained access to the dihaploid rose reference sequence of the *Rosa chinensis* variety 'Old Blush' (Hibrand *et al.,* 2018). The reference sequence is assembled to the level of pseudo-molecules representing all seven chromosomes and contains 90.1% to 96.1% of the whole genome sequence according to k-mer analysis. By using large sets of RNAseq data, among others the data presented in this thesis (chapter 6.2), 44,481 coding genes were predicted including 39,669 protein coding genes. This number is comparable to the amount of genes predicted with the Augustus tool in the 88 genome (Table 6.2). The BUSCO analysis of the predicted gene models of the 'Old Blush' genome indicates that only 3.4% of the orthologs are missing (Hibrand *et al.,* 2018), which is comparable to other fully sequenced genomes. Due to the higher quality of the 'Old Blush' genome and its gene prediction and the fact that it will be established as the rose reference sequence, this sequence was chosen for the analysis of the transcriptomic data and the work on the 88 genome was not followed up.

However, the draft genome of the genotype 88/124-46 is still a valuable resource, especially for the analysis of Rdr1 locus, which is not included in the genome of the variety 'Old Blush'. Together with other draft genomes like the published genomes of *Rosa multiflora* (Nakamura *et al.*, 2017) and *Rosa x damascena* (NCBI accession number LYNE00000000), the sequence of the 88 genome can be used for diversity studies or the analysis of gene and genome variations. In addition, the generated genome is the best source for primer development for the analysis of many of the populations used at the department, which derived from crossings with this genotype or relatives of this genotype.

7.5. Mechanisms of the rose defence response to fungal pathogens

One of the main goals of this thesis was to analyse transcriptomic changes in the rose leaf in response to fungal pathogens, in particular to the black spot fungus *D. rosae* (chapter 4).

The analysis of the compatible interaction of the susceptible rose variety PC with the *D. rosae* isolate DortE4 and a greenhouse mixture of *P. pannosa* revealed the up-regulation of many defence related mechanisms. The up-regulation of genes belonging to the class of PR10 and chitinases was similar in response to both pathogens. This might indicate that these mechanisms are part of a PTI-reaction triggered by an elicitor like e.g. fungal chitin. In addition, in response to *D. rosae* additional PRs genes and genes encoding or key enzymes of the lignin biosynthesis and the flavonoid pathway were up-regulated as well as genes of the SA-mediated signalling pathway. From unpublished results we know that the genome of PC contains genes which are similar to the *RGA1*-gene of the Rdr1-locus but do not mediate resistance. Furthermore, microscopic investigations indicate that a strong defence response occurs during the compatible interaction DortE4, but it seems that this reaction happens with temporal delay compared to a resistance reaction. This might indicate that the additional defence response to *D. rosae* is a partial ETI-reaction, which could be triggered by broken R-genes.

These findings could be the starting point of various experiments. The transcriptomic data of the incompatible interaction of the resistant rose genotype 91/100-5 with DortE4 presented in chapter 6.2 can give insights into the regulation of the presented genes and pathways during an incompatible interaction to estimate their importance for the plant defence in particular for the resistance reaction. Additionally, the reactions can be validated in another genetic background.

Further experiments should be performed to analyse the importance of the lignin biosynthesis and the flavonoid pathway. Determination of the lignin content e.g. by photometric measurement of leaf extracts (Schenk and Schikora, 2015) of infected and unaffected leaf tissue can clarify whether the lignin biosynthesis is really induced during the rose defence. Metabolomic analyses can give insights into the products of the flavonoid pathway which are synthesised in response to pathogenic attacks. So far the transcriptomic data indicate that the pathway leading to the synthesis of

flavonols is mainly affected in response to *D. rosae*. Especially genes encoding for flavonol synthases are strongly up-regulated. However, in *Fragaria* catechin, a proanthocyanidin, seems to be the most important product of this pathway in response to pathogens (Puhl and Treutter, 2008; Yamamoto *et al.*, 2000). Thus, it would be very interesting to see if the function of the flavonoid pathway is similar in the defence response of different rosacea species.

As mentioned before, the transcriptomic data might indicate that the defence response to *D. rosae* is comprised of a PTI and partial ETI reaction. Due to the fact that chitin is the main elicitor of the PTI reaction to fungi, it would be interesting to analyse the response of roses to an artificial application of chitin. Results of this experiment could help to distinguish between mechanisms which are more PTI or ETI-related. Furthermore, the generation of RNAi-plants of PC where the whole RGA gene family is down-regulated might prove if the additional response to *D. rosae* is ETI-related and mediated by members of this family.

One of the very surprising outcomes of the comparison of the response to the hemibiotrophic fungus *D. rosae* and the obligate biotrophic fungus *P. pannosa* was that PR1 and other marker genes of the SA-signalling pathway were up-regulated only in response to *D. rosae*. In the literature the SA-signalling pathway is described as one of the main pathways activating the defence response to biotrophic and hemibiotrophic pathogens (van Loon *et al.* 2006; Derksen *et al.*, 2013). It could be that the response of these genes to *P. pannosa* is simply not detected due to a lower inoculation density in the experiments with this pathogen. However, PR1, which was included in the set of genes tested for validating the MACE-data by qPCR, showed a significant down-regulation in response to *P. pannosa* in this experiment. This is a hint that the SA-pathway is only active in response to *D. rosae* and might even indicate that it is actively down-regulated in response to *P. pannosa,* but more genes need to be tested to confirm this. Nevertheless, the activation of the SA-signalling pathway could be one of the reasons activating the additional defence mechanisms in response to *D. rosae*. Taking all these findings together, the following model of the different defence responses to both pathogens can be hypothesised (Figure 7.1).

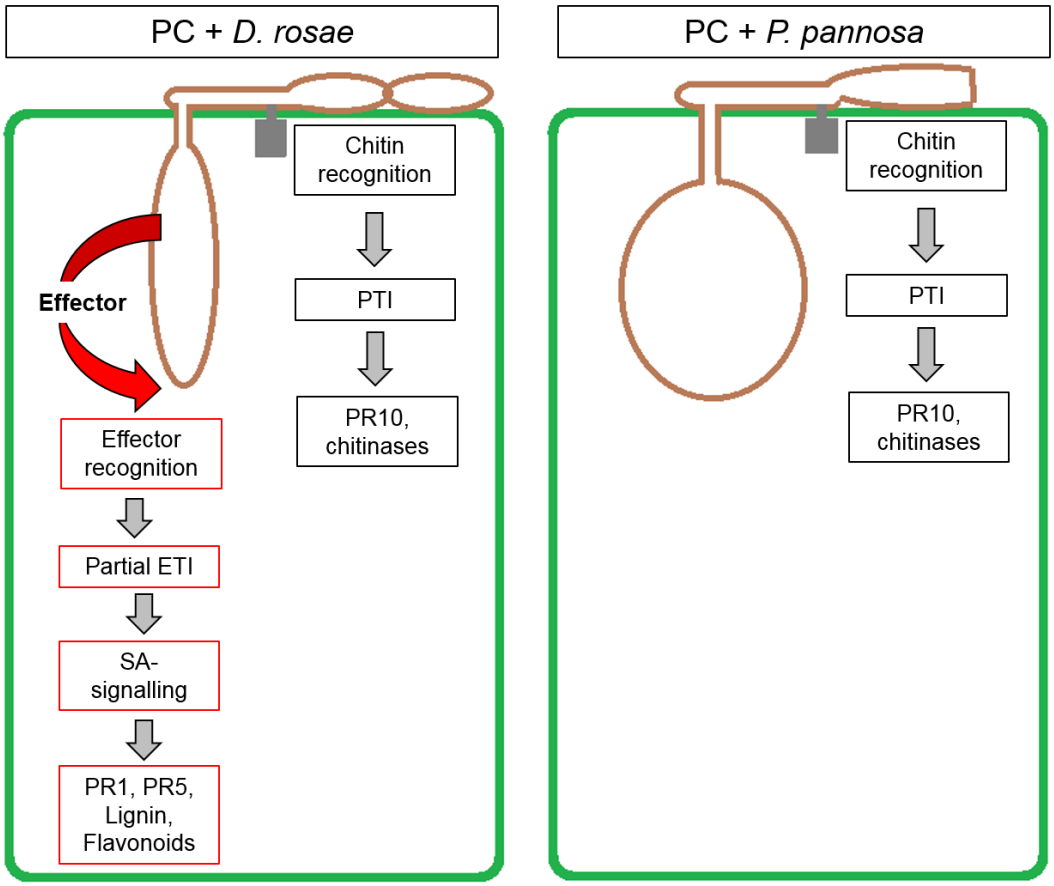


Figure 7.1: Hypothesised model of the defence response to *D. rosae* and *P. pannosa* during the compatible interaction with the rose variety PC. The recognition of chitin or another PAMP leads to the induction of a raid PTI reaction, which is characterised by the up-regulation of several homologs of PR10 and chitinases. This response is similar to both pathogens. Additionally, there is specific response to *D. rosae*. This response might be a partial ETI-reaction triggered by an effector recognised by a homolog of the *RGA1* gene of the Rdr1-locus. This recognition activates the SA-mediated signalling pathway which at the end leads to an up-regulation of additional PR-genes like PR1 and PR5 and genes of the lignin biosynthesis and the flavonoid pathway.

In addition to the general description of response mechanisms in the interaction with both pathogens, the transcriptomic analysis was performed to identify candidate genes for further analysis to investigate the influence of P-starvation on the defence response. For this purpose the hydroponic system was established (chapter 6.4). Besides the various response genes mentioned before (chapter 5), the genes encoding for WRKY75 are of particular interest for the analysis of the cross-talk between P-starvation and the plant defence. As shown here, two WRKY75 genes are

up-regulated in response to *D. rosae* and *P. pannosa*. Additionally, the importance of this WRKY transcription factor for the basal immune response and the resistance reaction was shown among others in *Fragaria* and Arabidopsis (Encinas-Villarejo *et al.*, 2009). However, before the influence of WRKY75 on the defence response was demonstrated it was described in Arabidopsis to be involved in regulating the response to P-starvation (Devaiah *et al.*, 2007). Suppression of this gene via RNAi resulted in plants which were more susceptible to P-stress and with a reduced P-uptake. Various genes for the response to P-starvation were down-regulated in roots of the RNAi-plants and the root architecture was influenced. These findings indicate that WRKY75 is involved in both the plant defence and the response to P-starvation, making it a promising candidate for the analysis of the cross-talk between both mechanisms.

7.6. Conclusion

This study aimed to generate crucial resources which are necessary for the analysis of the molecular mechanisms of the interaction between the rose and the black spot fungus *D. rosae*. Due to this fact, a draft genomic sequence of the *D. rosae* isolate DortE4 was sequenced. The analysis of the *D. rosae* genome revealed that a large proportion of the genome is duplicated, which might be an indication for a whole genome duplication. Additionally, the genome was used to predict the fungal secretome and identify potential effector genes which can be used to characterise functional vir and avr factors to gain crucial insight into the molecular mechanisms of the fungal infection. Furthermore, avr genes can be used to screen for additional R-genes and to identify the target in the plant, which can improve the breeding for more tolerant plant varieties. Besides the analysis of vir and avr factors, the sequence of *D. rosae* can be a tool to generate molecular markers for population studies, to clarify the phylogenetic position of *D. rosae* and its isolates or to identify the mating locus.

Besides the *D. rosae* genome, a draft sequence for the rose genotype 88/124-46 was generated but not further analysed because the rose reference sequence of the variety ‘Old Blush’ was completed in 2017. Nevertheless, the generated sequence is still a valuable resource and can be of use for further studies like diversity analysis or genome comparisons.

These resources were necessary to analyse the changes in the leaf transcriptome during the compatible interaction of the rose variety PC with *D. rosae* and *P. pannosa,* which indicates that there is a similar response to both pathogens in the form of an up-regulation of PR10-genes, chitinases and WRKY-transcription factors and pathogen-specific reactions. Of particular interest is the up-regulation of genes of the phenylpropanoid and flavonoid pathways as well as of the salicylic acid-signalling pathway including PR1 in response to *D. rosae*, which might indicate an additional defence response triggered by a partial ETI-reaction. Further analysis of the mechanisms described in this work can improve our knowledge of the defence response of roses to fungal pathogens.

References

Ade J., DeYoung B.J., Golstein C., Innes R.W. (2007) Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease. *Proceedings of the National Academy of Sciences of the Unites States of America*, 104(7), 2531–2536.

Ainsworth G.C. (2008) Ainsworth & Bisby's dictionary of the fungi. 10th ed. CAB International, Wallingford.

Albertin W., Marullo P. (2012) Polyploidy in fungi: Evolution after whole-genome duplication. *Proceedings of the Royal Society of London B: Biological Sciences*, 279, 2497–2509.

Amil-Ruiz F., Blanco-Portales R., Muñoz-Blanco J., Caballero J.L. (2011) The strawberry plant defense mechanism: A molecular review. *Plant & cell physiology*, 52(11), 1873–1903.

Amil-Ruiz F., Garrido-Gala J., Gadea J., Blanco-Portales R., Munoz-Merida A., Trelles O., los Santos B. de, Arroyo F.T., Aguado-Puig A., Romero F., Mercado J.-A., Pliego-Alfaro F., Munoz-Blanco J., Caballero J.L. (2016) Partial Activation of SA- and JA-Defensive Pathways in Strawberry upon Colletotrichum acutatum Interaction. *Frontiers in Plant Science*, 7, 1036.

Anderson R.G., Deb D., Fedkenheuer K., McDowell J.M. (2015) Recent Progress in RXLR Effector Research. Molecular Plant-Microbe Interactions, 28(10), 1063–1072.

Antolin-Llovera M., Ried M.K., Binder A., Parniske M. (2012) Receptor Kinase Signaling Pathways in Plant-Microbe Interactions. Annual Reviews of Phytopathology*,* 50, 451–473.

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*, 408(6814), 796–815.

Aronescu A. (1934) Diplocarpon Rosae: From Spore Germination to Haustorium Formation. *Bulletin of the Torrey Botanical Club*, 61(6), 291–329.

Ashburner M., Ball C.A., Blake J.A., Botstein D., Butler H., Cherry J.M., Davis A.P., Dolinski K., Dwight S.S., Eppig J.T., Harris M.A., Hill D.P., Issel-Tarver L., Kasarskis A., Lewis S., Matese J.C., Richardson J.E., Ringwald M., Rubin G.M., Sherlock G., Gene Ontology Consortium (2000) Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(1), 25–29.

Babalar M., Asghari M., Talaei A., Khosroshahi A. (2007) Effect of pre- and postharvest salicylic on ethylene production, fungal decay acid treatment and overall quality of Selva strawberry fruit. *Food Chemistry*, 105(2), 449–453.

Balan B., Marra F.P., Caruso T., Martinelli F. (2018) Transcriptomic responses to biotic stresses in Malus x domestica: A meta-analysis study. *Scientific reports*, 8(1), 1970.

Bent A.F., Mackey D. (2007) Elicitors, effectors, and R genes: The new paradigm and a lifetime supply of questions. *Annual Review of Phytopathology*, 45: 399–436.

Bhuiyan N.H., Selvaraj G., Wei Y., King J. (2009) Role of lignification in plant defense. *Plant Signaling & Behavior*, 4(2), 158–159.

Boatwright J.L., Pajerowska-Mukhtar K. (2013) Salicylic acid: An old hormone up to new tricks. *Molecular plant pathology*, 14(6), 623–634.

Brands S.J. (1989) The Taxonomicon. Universal Taxonomic Services, Zwaag, The Netherlands. http://taxonomicon.taxonomy.nl/. (Accessed 26.04.2018)

Brichet H. (2003) Distribution and Ecology: Overview. In: A. V. Roberts, T. Debener, S. Gudin (Eds). *Encyclopedia of rose science*. Elsevier Academic press, Oxford, 199.

Brisson L.F., Tenhaken R., Lamb C. (1994) Function of Oxidative Cross-Linking of Cell Wall Structural Proteins in Plant Disease Resistance. *The Plant Cell*, 6(12), 1703–1712.

Browne L.M., Conn K.L., Ayert W.A., Tewari J.P. (1991) The camalexins: New phytoalexins produced in the leaves of camelina sativa (cruciferae). *Tetrahedron*, 47(24), 3909–3914.

Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., Madden T.L. (2009) BLAST plus: Architecture and applications. *BMC Bioinformatics*, 10(1), 421.

Cantarel B.L., Korf I., Robb, Sofia M. C., Parra G., Ross E., Moore B., Holt C., Alvarado A.S., Yandell M. (2008) MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research*, 18(1), 188–196.

Casado-Díaz A., Encinas-Villarejo S., Santos B.d.l., Schilirò E., Yubero-Serrano E.-M., Amil-Ruíz F., Pocovi M.I., Pliego-Alfaro F., Dorado G., Rey M., Romero F., Muñoz-Blanco J., Caballero J.-L. (2006) Analysis of strawberry genes differentially

expressed in response to Colletotrichum infection. *Physiologia Plantarum*, 128(4), 633–650.

Catanzariti A.-M., Jones D.A. (2010) Effector proteins of extracellular fungal plant pathogens that trigger host resistance. *Functional Plant Biology*, 37(10), 901–906.

Chaanin A. (2003) Selection strategies for cut roses. In: A. V. Roberts, T. Debener, S. Gudin (Eds). *Encyclopedia of rose science*. Elsevier Academic press, Oxford, 33-41.

Chevalier M., Parisi L., Gueye B., Campion C., Simoneau P., Poupard P. (2008) Specific activation of PR-10 pathogenesis-related genes in apple by an incompatible race of Venturia inaequalis. *Biologia Plantarum*, 52(4), 718–722.

Conesa A., Gotz S., Garcia-Gomez J.M., Terol J., Talon M., Robles M. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, 21(18), 3674–3676.

Conesa A., Madrigal P., Tarazona S., Gomez-Cabrero D., Cervera A., McPherson A., Szcześniak M.W., Gaffney D.J., Elo L.L., Zhang X., Mortazavi A. (2016) A survey of best practices for RNA-seq data analysis. *Genome Biology*, 17(1), 13.

Cova V., Paris R., Toller C., Patocchi A., Velasco R., Komjanc M. (2017) Apple genes involved in the response to Venturia inaequalis and salicylic acid treatment. *Scientia Horticulturae*, 226, 157–172.

Cui H., Tsuda K., Parker J.E. (2015) Effector-triggered immunity: from pathogen perception to robust defense. *Annual Review of Plant Biology*, 66, 487–511.

Cuomo C.A., Güldener U., Xu J.-R., Trail F., Turgeon B.G., Di Pietro A., Walton J.D., Ma L.-J., Baker S.E., Rep M., Adam G., Antoniw J., Baldwin T., Calvo S., Chang Y.-L., DeCaprio D., Gale L.R., Gnerre S., Goswami R.S., Hammond-Kosack K., Harris L.J., Hilburn K., Kennell J.C., Kroken S., Magnuson J.K., Mannhaupt G., Mauceli E., Mewes H.-W., Mitterbauer R., Muehlbauer G., Münsterkötter M., Nelson D., O'donnell K., Ouellet T., Qi W., Quesneville H., Roncero M.I.G., Seong K.-Y., Tetko I.V., Urban M., Waalwijk C., Ward T.J., Yao J., Birren B.W., Kistler H.C. (2007) The Fusarium graminearum genome reveals a link between localized polymorphism and pathogen specialization. *Science*, 317(5843), 1400–1402.

Debener T. (1999) Genetic Analysis of Horticulturally Important Morphological and Physiological Characters in Diploid Roses. *Gartenbauwissenschaft*, 64(1), 14–20.

Debener T., Byrne D.H. (2014) Disease resistance breeding in rose: Current status and potential of biotechnological tools. *Plant Science*, 228, 107–117.

Debener T., Linde M. (2009) Exploring Complex Ornamental Genomes: The Rose as a Model Plant. *Critical Reviews in Plant Sciences*, 28(4), 267–280.

Derksen H., Rampitsch C., Daayf F. (2013) Signaling cross-talk in plant disease resistance. *Plant Science*, 207, 79–87.

Devaiah B.N., Karthikeyan A.S., Raghothama K.G. (2007) WRKY75 transcription factor is a modulator of phosphate acquisition and root development in Arabidopsis. *Plant Physiology*, 143(4), 1789–1801.

DeYoung B.J., Innes R.W. (2006) Plant NBS-LRR proteins in pathogen sensing and host defense. *Nature Immunology*, 7(12), 1243–1249.

Drewes-Alvarez R. (1992) Untersuchungen am Pathosystem Sternrusstau (Marssonina rosae [Lib.] Died.)-Rose (Rosa L.): zur Klärung der Frage nach pilzlichen Rassen und zur Übertragung der Resistenz aus der diploiden Rosa multiflora Thunb. auf tetraploide Gartenrosen. Dissertation. Universität Hamburg

Drewes-Alvarez R. (2003) Black Spot. In: A. V. Roberts, T. Debener, S. Gudin (Eds). *Encyclopedia of rose science*. Elsevier Academic press, Oxford: 148–153.

Dugo M.L., Satovic Z., Millán T., Cubero J.I., Rubiales D., Cabrera A., Torres A.M. (2005) Genetic mapping of QTLs controlling horticultural traits in diploid roses. *Theoretical and Applied Genetics*, 111(3), 511–520.

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., Kuees 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, Yves, Rouze P., Ellis J.G., Dodds P.N., Schein J.E., Zhong S., Hamelin R.C., Grigoriev I.V., Szabo L.J., Martin F. (2011) Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proceedings of the National Academy of Sciences of the Unites States of America*, 108(22), 9166–9171.

Ellinger D., Naumann M., Falter C., Zwikowics C., Jamrow T., Manisseri C., Somerville S.C., Voigt C.A. (2013) Elevated early callose deposition results in complete penetration resistance to powdery mildew in Arabidopsis. *Plant Physiology*, 161(3), 1433–1444.

Encinas-Villarejo S., Maldonado A.M., Amil-Ruiz F., los Santos B. de, Romero F., Pliego-Alfaro F., Muñoz-Blanco J., Caballero J.L. (2009) Evidence for a positive regulatory role of strawberry (Fragaria x ananassa) Fa WRKY1 and Arabidopsis At WRKY75 proteins in resistance. *Journal of Experimental Botany*, 60(11), 3043–3065.

Feys B.J., Moisan L.J., Newman M.A., Parker J.E. (2001) Direct interaction between the Arabidopsis disease resistance signaling proteins, EDS1 and PAD4. *The EMBO journal*, 20(19), 5400–5411.

Flor H.H. (1971) Current status of the gene-for-gene concept. *Annual Review of Phytopathology*, 9(1), 275-296.

FloraHolland (2016) https://www.royalfloraholland.com/en (Accessed 26.04.2018)

Frick L. (1943) Untersuchungen über Biologie und Pathogenität von Diplocarpon rosae (Lib.) Wolf. Dissertation, Eidgenössischen Technischen Hochschule, Zürich.

Gachomo E.W., Dehne H.-W., Steiner U. (2006) Microscopic evidence for the hemibiotrophic nature of Diplocarpon rosae, cause of black spot disease of rose. *Physiological and Molecular Plant Pathology*, 69(1-3), 86–92.

GenXPro (2015) See why. Massive Analysis of cDNA Ends (MACE). Available from genxpro.net/wp-content/uploads/2015/07/MACE_GenXPro.pdf (Accessed 26.04.2018).

Girard V., Dieryckx C., Job C., Job D. (2013) Secretomes: The fungal strike force. *Proteomics*, 13(3-4), 597–608.

Godfrey D., Bohlenius H., Pedersen C., Zhang Z., Emmersen J., Thordal-Christensen H. (2010) Powdery mildew fungal effector candidates share N-terminal Y/F/WxC-motif. *BMC Genomics*, 11(1), 317.

Gomez-Gomez L., Boller T. (2000) FLS2: An LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. *Molecular Cell*, 5(6), 1003–1011.

Gudin S. (2000) Rose: Genetics and breeding. *Plant Breeding Reviews*, 17, 159–189.

Gudin S. (2003) Breeding: Overview. In: A. V. Roberts, T. Debener, S. Gudin (Eds). *Encyclopedia of rose science*. Elsevier Academic press, Oxford, 25-31.

Guidarelli M., Carbone F., Mourgues F., Perrotta G., Rosati C., Bertolini P., Baraldi E. (2011) Colletotrichum acutatum interactions with unripe and ripe strawberry fruits

and differential responses at histological and transcriptional levels. *Plant Pathology*, 60(4), 685–697.

Hattendorf A., Linde M., Mattiesch L., Debener T., Kaufmann H. (2004) Genetic analysis of rose resistance genes and their localisation in the rose genome. *Acta Horticulturae*, 651, 123–130.

Heinrichs, F. (2008) International Statistics Flowers and Plants, Vol. 56. AIPH, Union Fleurs, Brussels, Belgium.

Hibrand L., Ruttink T., Hamama L., Kirov I., Lakhwani D., Zhou N.-N., Bourke P., Daccord N., Leus L., Schulz D., van deGeest H., Hesselink T., van Laere K., Balzergue S., Thouroude T., Chastellier A., Jeauffre J., Voisine L., Gaillard S., Borm T., Arens P., Voorrips R., Maliepaard C., Neu E., Linde M., Le Paslier M.-C., Berard A., Bounon R., Clotault J., Choisne N., Quesneville H., Kawamura K., Aubourg S., Sakr S., Smulder R., Schijlen E., Bucher E., Debener T., Riek J. de, Foucher F. (2018) A high-quality sequence of Rosa chinensis to elucidate genome structure and ornamental traits. *Nature Plants*, 4, 473-484.

Hirai N., Sugie M., Wada M., Lahlou E.H., Kamo T., Yoshida R., Tsuda M., Ohigashi H. (2000) Triterpene phytoalexins from strawberry fruit. *Bioscience, Biotechnology, and Biochemistry*, 64(8), 1707–1712.

Horst R.K., Cloyd R.A. (2007) *Rose diseases and pests,* 2nd ed. APS Press, St. Paul.

Hosseini Moghaddam H., Leus L., Riek J., Huylenbroeck J., Bockstaele E. (2012) Construction of a genetic linkage map with SSR, AFLP and morphological markers to locate QTLs controlling pathotype-specific powdery mildew resistance in diploid roses. *Euphytica*, 184(3), 413–427.

Hu M., Polyak K. (2006) Serial Analysis of Gene Expression. *Nature Protocols*, 1(4), 1743–1760.

Hubert D., Römhelm V., Weinmann M. (2012) Relationship between Nutrition, Plant Diseases and Pests. In: H. Marschner, P. Marschner (Eds). *Marschner's Mineral nutrition of higher plants*, 3rd ed. Elsevier Academic press, Amsterdam: 283–298.

Iakimova, E. T., Michalczuk, L., & Woltering, E. J. (2005) Hypersensitive cell death in plants-its mechnisms and role in plant defence against pathogens. *Journal of Fruit and Ornamental Plant Research*, 13(1-4), 135–158.

Jiang C., Cheng X., Cheng Q., Huang M., Xu L. (2014) High-efficiency transformation of the plant pathogenic fungus Marssonina brunnea. *Journal of Plant Pathology*, 9683), 459–464.

Jiao W.-B., Schneeberger K. (2017) The impact of third generation genomic technologies on plant genome assembly. *Current Opinion in Plant Biology*, 36, 64–70.

Jones, Jonathan D. G., Dangl J.L. (2006) The plant immune system. *Nature* , 444(7117), 323–329.

Jonge R. de, van Esse, H. Peter, Kombrink A., Shinya T., Desaki Y., Bours R., van der Krol, Sander, Shibuya N., Joosten, Matthieu H. A. J., Thomma, Bart P. H. J. (2010) Conserved Fungal LysM Effector Ecp6 Prevents Chitin-Triggered Immunity in Plants. *Science*, 329(5994), 953–955.

Kaemper J., Kahmann R., Boelker M., Ma L.-J., Brefort T., Saville B.J., Banuett F., Kronstad J.W., Gold S.E., Mueller O., Perlin M.H., Woesten H.A.B., Vries R. de, Ruiz-Herrera J., Reynaga-Pena C.G., Snetselaar K., McCann M., Perez-Martin J., Feldbruegge 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., Muench K., Roessel 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., Haeuser-Hahn I., Vaupel M., Koopmann E., Friedrich G., Voss H., Schlueter T., Margolis J., Platt D., Swimmer C., Gnirke A., Chen F., Vysotskaia V., Mannhaupt G., Gueldener U., Muensterkoetter 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., Birren B.W. (2006) Insights from the genome of the biotrophic fungal plant pathogen Ustilago maydis. *Nature*, 444 (7115), 97–101.

Kahl G., Molina C., Rotter B., Jüngling R., Frank A., Krezdorn N., Hoffmeier K., Winter P. (2012) Reduced representation sequencing of plant stress transcriptomes*. Journal of Plant Biochemistry and Biotechnology*, 21(1), 119–127.

Kaufmann H., Mattiesch L., Lörz H., Debener T. (2003) Construction of a BAC library of Rosa rugose Thunb. and assembly of a contig spanning Rdr1, a gene that confers resistance to blackspot. *Molecular Genetics and Genomics*, 268(5), 666–674.

Kaufmann H., Qiu X., Wehmeyer J., Debener T. (2012) Isolation, Molecular Characterization, and Mapping of Four Rose MLO Orthologs. *Frontiers in Plant Science*, 3(244).

Kazan K., Manners J.M. (2013) MYC2: The Master in Action. *Molecular Plant*, 6(3), 686–703.

Kobayashi, N., Horikoshi, T., Katsuyama, H., Handa, T. and Takayanagi, K. (1998)A simple and efficient DNA extraction method for plants, especially woody plants. *Plant Tissue Culture and Biotechnology*, 4, 76–80.

Koning-Boucoiran, Carole F. S., Esselink G.D., Vukosavljev M., van 't Westende, Wendy P. C., Gitonga V.W., Krens F.A., Voorripsi R.E., van de Weg, W. Eric, Schulz D., Debener T., Maliepaardl C., Arens P., Smulders, Marinus J. M. (2015) Using RNA-Seq to assemble a rose transcriptome with more than 13,000 full-length expressed genes and to develop the WagRhSNP 68k Axiom SNP array for rose (Rosa L.). *Frontiers in Plant Science*, 6(249).

Korf I. (2004) Gene finding in novel genomes. *BMC Bioinformatics*, 5(59).

Kubicek C.P., Starr T.L., Glass N.L. (2014) Plant Cell Wall–Degrading Enzymes and Their Secretion in Plant-Pathogenic Fungi. *Annual Review of Phytopathology*, 52, 427–451.

Kullman B., Teterin W. (2006) Estimation of fungal geome size: comparison of image cytometry and photometric cytometry. *Folia Cryptogamica Estonica*, 2006, 42, 43–56.

Kunze G., Zipfel C., Robatzek S., Niehaus K., Boller T., Felix G. (2004) The N terminus of bacterial elongation factor Tu elicits innate immunity in Arabidopsis plants. *Plant Cell*, 16(12), 3496–3507.

Langcake P., Pryce R.J. (1976) The production of resveratrol by Vitis vinifera and other members of the Vitaceae as a response to infection or injury. *Physiological Plant Pathology*, 9(1), 77–86.

Lara A.J., Perez-Trabado G., Villalobos D.P., Diaz-Moreno S., Canton F.R., Claros M.G. (2007) A web tool to discover full-length sequences - Full-Lengther. In: *Innovations in Hybrid Intelligent Systems.* Springer, Berlin, Heidelberg, 361-368.

Laurie-Berry N., Joardar V., Street I.H., Kunkel B.N. (2006) The Arabidopsis thaliana JASMONATE INSENSITIVE 1 Gene Is Required for Suppression of Salicylic Acid-Dependent Defenses During Infection by Pseudomonas syringae. *Molecular Plant-Microbe Interactions*, 19(7), 789–800.

Lee S.-J., Rose J.K.C. (2012) A yeast secretion trap assay for identification of secreted proteins from eukaryotic phytopathogens and their plant hosts. *Methods in Molecular Biology*, 835, 519–530.

Li C., Lin F., An D., Wang W., Huang R. (2017a) Genome Sequencing and Assembly by Long Reads in Plants. *Genes*, 9(1), 6.

Li D., Tang Y., Lin J., Cai W. (2017b) Methods for genetic transformation of filamentous fungi. *Microbial Cell Factories*, 16(1), 168.

Linde M., Hattendorf A., Kaufmann H., Debener T. (2006) Powdery mildew resistance in roses: QTL mapping in different environments using selective genotyping. *Theoretical and Applied Genetics*, 113(6), 1081–1092.

Linde M., Mattiesch L., Debener T. (2004) Rpp1, a dominant gene providing race-specific resistance to rose powdery mildew (Podosphaera pannosa): Molecular mapping, SCAR development and confirmation of disease resistance data. *Theoretical and Applied Genetics*, 109(6), 1261–1266.

Liu X., Bai X., Wang X., Chu C. (2007) OsWRKY71, a rice transcription factor, is involved in rice defense response. *Journal of Plant Physiology*, 164(8), 969–979.

Liu Y., Zhang S. (2004) Phosphorylation of 1-Aminocyclopropane-1-Carboxylic Acid Synthase by MPK6, a Stress-Responsive Mitogen-Activated Protein Kinase, Induces Ethylene Biosynthesis in Arabidopsis. *The Plant Cell*, 16(12), 3386–3399.

Lo Presti L., Lanver D., Schweizer G., Tanaka S., Liang L., Tollot M., Zuccaro A., Reissmann S., Kahmann R. (2015) Fungal Effectors and Plant Susceptibility. *Annual Review of Plant Biology*, 66, 513–545.

Lomsadze A., Ter-Hovhannisyan V., Chernoff Y.O., Borodovsky M. (2005) Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic AcidsResearch*, 33(20), 6494–6506.

Lorenzo O., Chico J.M., Sanchez-Serrano J.J., Solano R. (2004) Jasmonate-insensitive1 encodes a MYC transcription factor essential to discriminate between different jasmonate-regulated defense responses in Arabidopsis. *Plant Cell*, 16(7), 1938–1950.

Lorenzo O., Piqueras R., Sanchez-Serrano J.J., Solano R. (2003) ETHYLENE RESPONSE FACTOR1 integrates signals from ethylene and jasmonate pathways in plant defense. *Plant Cell*, 15(1), 165–178.

Lowe T.M., Eddy S.R. (1997) tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research*, 25(5), 955–964.

Lu Y., Chen Q., Bu Y., Luo R., Hao S., Zhang J., Tian J., Yao Y. (2017) Flavonoid Accumulation Plays an Important Role in the Rust Resistance of Malus Plant Leaves. *Frontiers in Plant Science*, 8, 1286.

Luderer R., Takken F.L.W., Wit P. de, Joosten M. (2002) Cladosporium fulvum overcomes Cf-2-mediated resistance by producing truncated AVR2 elicitor proteins. *Molecular Microbiology*, 45(3), 875–884.

Lühmann A.-K. (2010) Genetische Diversität des Sternrußtaus (Diplocarpon rosae) an Rosen. Dissertation, Leibniz Universität, Hannover.

Ma L., Lukasik E., Gawehns F., Takken F.L.W. (2012) The use of agroinfiltration for transient expression of plant resistance and fungal effector proteins in Nicotiana benthamiana leaves. *Methods in Molecular Biology,* 835, 61–74.

Ma L.-J., Fedorova N.D. (2010) A practical guide to fungal genome projects: Strategy, technology, cost and completion. *Mycology*, 1(1), 9–24.

Ma L.-J., Ibrahim A.S., Skory C., Grabherr M.G., Burger G., Butler M., Elias M., Idnurm A., Lang B.F., Sone T., Abe A., Calvo S.E., Corrochano L.M., Engels R., Fu J., Hansberg W., Kim J.-M., Kodira C.D., Koehrsen M.J., Liu B., Miranda-Saavedra D., O'Leary S., Ortiz-Castellanos L., Poulter R., Rodriguez-Romero J., Ruiz-Herrera J., Shen Y.-Q., Zeng Q., Galagan J., Birren B.W., Cuomo C.A., Wickes B.L. (2009) Genomic analysis of the basal lineage fungus Rhizopus oryzae reveals a whole-genome duplication. *PLoS Genetics*, 5(7), e1000549.

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.,

Rep M. (2010) Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. *Nature*, 464(7287), 367–373.

Malek B. von, Debener T. (1998) Genetic analysis of resistance to blackspot (Diplocarpon rosae) in tetraploid roses. *Theoretical and Applied Genetics*, 96(2), 228–231.

Malone J.H., Oliver B. (2011) Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biology*, 9(1), 34.

Mao G., Meng X., Liu Y., Zheng Z., Chen Z., Zhang S. (2011) Phosphorylation of a WRKY transcription factor by two pathogen-responsive MAPKs drives phytoalexin biosynthesis in Arabidopsis. *The Plant Cell*, 23(4), 1639–1653.

Meinke D.W. (1998) Arabidopsis thaliana: A Model Plant for Genome Analysis. *Science*, 282(5389), 662–682.

Meinken J., Asch D.K., Neizer-Ashun K.A., Chang G.-H., R.Cooper JR C., Min X.J. (2014) FunSecKB2: a fungal protein subcellular location knowledgebase. *Computational Molecular Biology*, 4(4).

Menz I., Straube J., Linde M., Debener T. (2017) The TNL gene Rdr1 confers broad-spectrum resistance to Diplocarpon rosae. *Molecular Plant Pathology*, 19(5), 1104-1113.

Michael T.P., VanBuren R. (2015) Progress, challenges and the future of crop genomes. *Current Opinion in Plant Biology*, 24, 71–81.

Milholland R.D. (1982) Histopathology of Strawberry Infected with Colletotrichum fragariae. *Phytopathology*, 72(11), 1434.

Mohanta T.K., Bae H. (2015) The diversity of fungal genome. *Biological Procedures*, 17(1), 8.

Nakamura N., Hirakawa H., Sato S., Otagaki S., Matsumoto S., Tabata S., Tanaka Y. (2017) Genome structure of Rosa multiflora, a wild ancestor of cultivated roses. *DNA Research*, 25(2), 113-121

Nürnberger T., Brunner F., Kemmerling B., Piater L. (2004) Innate immunity in plants and animals: striking similarities and obvious differences. *Immunological Reviews*, 198(1), 249–266.

Nybom H., Werlemark G. (2017) Realizing the Potential of Health-Promoting Rosehips from Dogroses (Rosa sect. Caninae). *Current Bioactive Compounds*, 13(1), 3–17.

Blechert O. (2005) Molekularbiologische und mikroskopische Untersuchungen am Pathosystem Diplocarpon rosae Wolf (Sternrusstau) - Rose. Dissertation, Leibniz Universität, Hannover.

Pajerowska-Mukhtar K.M., Emerine D.K., Mukhtar M.S. (2013) Tell me more: roles of NPRs in plant immunity. *Trends in Plant Science*, 18(7), 402–411.

Peng M., Kuc J. (1992) Peroxidase-generated hydrogen peroxide as a source of antifungal activity in vitro and on tobacco leaf disks. *Phytopathology*, 82(6), 696–699.

Piasecka A., Jedrzejczak-Rey N., Bednarek P. (2015) Secondary metabolites in plant innate immunity: Conserved function of divergent chemicals. *The New Phytologist*, 206(3), 948–964.

Pieterse, Corne M. J., Leon-Reyes A., Van der Ent, Sjoerd, Van Wees, Saskia C. M. (2009) Networking by small-molecule hormones in plant immunity. *Nature Chemical Biology*, 5(5), 308–316.

Pontier D., Balagué C., Roby D. (1998) The hypersensitive response. A programmed cell death associated with plant resistance. *Comptes Rendus de l'Académie des Sciences - Series III - Sciences de la Vie*, 321(9), 721–734.

Potter D., Eriksson T., Evans R.C., Oh S., Smedmark, J. E. E., Morgan D.R., Kerr M., Robertson K.R., Arsenault M., Dickinson T.A., Campbell C.S. (2007) Phylogeny and classification of Rosaceae. *Plant Systematics and Evolution*, 266(1-2), 5–43.

Poupard P., Parisi L., Campion C., Ziadi S., Simoneau P. (2003) A wound- and ethephon-inducible PR-10 gene subclass from apple is differentially expressed during infection with a compatible and an incompatible race of Venturia inaequalis. *Physiological and Molecular Plant Pathology*, 62(1), 3–12.

Prabhu A.S., Fageria N.K., Berni R.F. (2007) Phosphorus and Plant disease. In: L. E. Datnoff (Ed). *Mineral nutrition and plant disease*. APS Press American Phytopathological Society, St. Paul, USA, 45–55.

Puhl I., Treutter D. (2008) Ontogenetic variation of catechin biosynthesis as basis for infection and quiescence of Botrytis cinerea in developing strawberry fruits. *Journal of Plant Diseases and Protection*, 115(6), 247–251.

Pühringer H., Moll D., Hoffmann-Sommergruber K., Watillon B., Katinger H., da Câmara Machado M.L. (2000) The promoter of an apple Ypr10 gene, encoding the major allergen Mal d 1, is stress- and pathogen-inducible. *Plant Science*, 152, 35–50.

Raffaele S., Kamoun S. (2012) Genome evolution in filamentous plant pathogens: Why bigger can be better. *Nature Reviews. Microbiology*, 10(6), 417–430.

Roberts A.V., Debener T., Gudin S. (2003) Introduction. In: A. V. Roberts, T. Debener, S. Gudin (Eds). *Encyclopedia of rose science*. Elsevier Academic press, Oxford, VI-VII.

Rohe M., Gierlich A., Hermann H., Hahn M., Schmidt B., Rosahl S., Knogge W. (1995) The race-specific elicitor, NIP1, from the barley pathogen, Rhynchosporium secalis, determines avirulence on host plants of the Rrs1 resistance genotype. *The EMBO Journal*, 14(17), 4168–4177.

Rooney H.C.E., van 't Klooster, JW, van der Hoorn, RAL, Joosten M., Jones J.D.G., Wit P. de (2005) Cladosporium Avr2 inhibits tomato Rcr3 protease required for Cf-2-dependent disease resistance. *Science*, 308(5729), 1783–1786.

Safár J., Simková H., Kubaláková M., Číhalíková J., Suchánková P., Bartos J., Dolezel J. (2010) Development of chromosome-specific BAC resources for genomics of bread wheat. *Cytogenetic and Genome Research*, 129(1-3), 211–223.

Salazar S.M., Castagnaro A.P., Arias M.E., Chalfoun N., Tonello U., Díaz Ricci J.C. (2007) Induction of a defense response in strawberry mediated by an avirulent strain of Colletotrichum. *European Journal of Plant Pathology*, 117(2), 109–122.

Saunders D.G.O., Win J., Cano L.M., Szabo L.J., Kamoun S., Raffaele S. (2012) Using hierarchical clustering of secreted protein families to classify and rank candidate effectors of rust fungi. *PloS One*, 7(1), e29847.

Schatz M.C., Witkowski J., McCombie W.R. (2012) Current challenges in de novo plant genome sequencing and assembly. *Genome Biology*, 13(4), 243.

Schenk S., Schikora A. (2015) Lignin Extraction and Quantification, a Tool to Monitor Defense Reaction at the Plant Cell Wall Level. *BIO-Protocol*, 5, e1430.

Schmidt S.M., Panstruga R. (2011) Pathogenomics of fungal plant parasites: What have we learnt about pathogenesis? *Current Opinion in Plant Biology*, 14(4), 392–399.

Selin C., de Kievit, Teresa R., Belmonte M.F., Fernando, W. G. Dilantha (2016) Elucidating the Role of Effectors in Plant-Fungal Interactions: Progress and Challenges. *Frontiers in Microbiology*, 7, 600.

Simao F.A., Waterhouse R.M., Ioannidis P., Kriventseva E.V., Zdobnov E.M. (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31(19), 3210–3212.

Slusarenko A.J., Fraser R.S.S., van Loon L.C. (Eds) (2000) *Mechanisms of resistance to plant diseases*. Kluwer, Dordrecht: 620 pp.

Spanu P.D., Abbott J.C., Amselem J., Burgis T.A., Soanes D.M., Stueber K., van Themaat E.V.L., Brown J.K.M., Butcher S.A., Gurr S.J., Lebrun M.-H., Ridout C.J., Schulze-Lefert P., Talbot N.J., Ahmadinejad N., Ametz C., Barton G.R., Benjdia M., Bidzinski P., Bindschedler L.V., Both M., Brewer M.T., Cadle-Davidson L., Cadle-Davidson M.M., Collemare J., Cramer R., Frenkel O., Godfrey D., Harriman J., Hoede C., King B.C., Klages S., Kleemann J., Knoll D., Koti P.S., Kreplak J., Lopez-Ruiz F.J., Lu X., Maekawa T., Mahanil S., Micali C., Milgroom M.G., Montana G., Noir S., O'Connell R.J., Oberhaensli S., Parlange F., Pedersen C., Quesneville H., Reinhardt R., Rott M., Sacristan S., Schmidt S.M., Schoen M., Skamnioti P., Sommer H., Stephens A., Takahara H., Thordal-Christensen H., Vigouroux M., Wessling R., Wicker T., Panstruga R. (2010) Genome Expansion and Gene Loss in Powdery Mildew Fungi Reveal Tradeoffs in Extreme Parasitism. *Science*, 330(6010), 1543–1546.

Sperschneider J., Dodds P.N., Gardiner D.M., Manners J.M., Singh K.B., Taylor J.M. (2015a) Advances and Challenges in Computational Prediction of Effectors from Plant Pathogenic Fungi. *PLoS Pathogens*, 11(5), e1004806.

Sperschneider J., Williams A.H., Hane J.K., Singh K.B., Taylor J.M. (2015b) Evaluation of Secretion Prediction Highlights Differing Approaches Needed for Oomycete and Fungal Effectors. *Frontiers in Plant Science*, 6, 1168.

Sperschneider J., Gardiner D.M., Dodds P.N., Tini F., Covarelli L., Singh K.B., Manners J.M., Taylor J.M. (2016) EffectorP: predicting fungal effector proteins from secretomes using machine learning. *New Phytologist*, 210(2), 743–761.

Spoel S.H., Koornneef A., Claessens S.M.C., Korzelius J.P., van Pelt J.A., Mueller M.J., Buchala A.J., Metraux J.P., Brown R., Kazan K., van Loon L.C., Dong X.N., Pieterse C.M.J. (2003) NPR1 modulates cross-talk between salicylate- and jasmonate-dependent defense pathways through a novel function in the cytosol. *Plant Cell*, 15(3), 760–770.

Spurgeon S.L., Jones R.C., Ramakrishnan R. (2008) High throughput gene expression measurement with real time PCR in a microfluidic dynamic array. *PloS one*, 3(2), e1662.

Stanke M., Keller O., Gunduz I., Hayes A., Waack S., Morgenstern B. (2006) AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Research*, 34, W435-W439.

Stankova H., Hastie A.R., Chan S., Vrana J., Tulpova Z., Kubalakova M., Visendi P., Hayashi S., Luo M., Batley J., Edwards D., Dolezel J., Simkova H. (2016) BioNano genome mapping of individual chromosomes supports physical mapping and sequence assembly in complex plant genomes. *Plant Biotechnology Journal*, 14(7), 1523–1531.

Stefanato F.L., Abou-Mansour E., Buchala A., Kretschmer M., Mosbach A., Hahn M., Bochet C.G., Métraux J.-P., Schoonbeek H.-j. (2009) The ABC transporter BcatrB from Botrytis cinerea exports camalexin and is a virulence factor on Arabidopsis thaliana. *The Plant Journal*, 58(3), 499–510.

Svetaz L.A., Bustamante C.A., Goldy C., Rivero N., Müller G.L., Valentini G.H., Fernie A.R., Drincovich M.F., Lara M.V. (2017) Unravelling early events in the Taphrina deformans-Prunus persica interaction: An insight into the differential responses in resistant and susceptible genotypes. *Plant, Cell and Environment*, 40(8), 1456–1473.

Tennessen J.A., Govindarajulu R., Ashman T.-L., Liston A. (2014) Evolutionary Origins and Dynamics of Octoploid Strawberry Subgenomes Revealed by Dense Targeted Capture Linkage Maps. *Genome Biology and Evolution*, 6(12), 3295–3313.

Terefe-Ayana D., Yasmin A., Le T.L., Kaufmann H., Biber A., Kühr A., Linde M., Debener T. (2011) Mining disease-resistance genes in roses: Functional and molecular characterization of the rdr1 locus. *Frontiers in Plant Science*, 2, 35.

Toljamo A., Blande D., Kärenlampi S., Kokko H. (2016) Reprogramming of Strawberry (Fragaria vesca) Root Transcriptome in Response to Phytophthora cactorum. *PloS one*, 11, e0161078.

Ton J., Mauch-Mani B. (2004) Beta-amino-butyric acid-induced resistance against necrotrophic pathogens is based on ABA-dependent priming for callose. *The Plant Journal*, 38(1), 119–130.

Torres M.A., Jones J.D.G., Dangl J.L. (2006) Reactive oxygen species signaling in response to pathogens. *Plant Physiology*, 141(2), 373–378.

Tsuda K., Katagiri F. (2010) Comparing signaling mechanisms engaged in pattern-triggered and effector-triggered immunity. *Current Opinion in Plant Biology*, 13(4), 459–465.

Tsuda K., Somssich I.E. (2015) Transcriptional networks in plant immunity. *New Phytologist*, 206(3), 932–947.

Ulker B., Somssich I.E. (2004) WRKY transcription factors: From DNA binding towards biological function. *Current Opinion in Plant Biology*, 7(5), 491–498.

van den Burg, Harrold A., Harrison S.J., Joosten, Matthieu H. A. J., Vervoort J., de Wit, Pierre J. G. M. (2006) Cladosporium fulvum Avr4 protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection. *Molecular Plant-Microbe Interaction*, 19(12), 1420–1430.

Van Der Biezen E.A., Jones J.D.G. (1998) Plant disease-resistance proteins and the gene-for-gene concept. *Trends in Biochemical Sciences*, 23(12), 454–456.

van Loon L.C., Rep M., Pieterse C.M.J. (2006) Significance of inducible defense-related proteins in infected plants. *Annual Review of Phytopathology*, 44, 135–162.

van't Slot, Klaas A. E., Gierlich A., Knogge W. (2007) A single binding site mediates resistance- and disease-associated activities of the effector protein NIP1 from the barley pathogen Rhynchosporium secalis. *Plant Physiology*, 144(3), 1654–1666.

Vela-Corcía D., Bautista R., Vicente A. de, Spanu P.D., Pérez-García A. (2016) De novo Analysis of the Epiphytic Transcriptome of the Cucurbit Powdery Mildew Fungus Podosphaera xanthii and Identification of Candidate Secreted Effector Proteins. *PloS one*, 11(10), e0163379.

Velasco R., Zharkikh A., Affourtit J., Dhingra A., Cestaro A., Kalyanaraman A., Fontana P., Bhatnagar S.K., Troggio M., Pruss D., Salvi S., Pindo M., Baldi P., Castelletti S., Cavaiuolo M., Coppola G., Costa F., Cova V., Dal Ri A., Goremykin V., Komjanc M., Longhi S., Magnago P., Malacarne G., Malnoy M., Micheletti D., Moretto M., Perazzolli M., Si-Ammour A., Vezzulli S., Zini E., Eldredge G., Fitzgerald L.M., Gutin N., Lanchbury J., Macalma T., Mitchell J.T., Reid J., Wardell B., Kodira C., Chen Z., Desany B., Niazi F., Palmer M., Koepke T., Jiwan D., Schaeffer S., Krishnan V., Wu C., Chu V.T., King S.T., Vick J., Tao Q., Mraz A., Stormo A., Stormo K., Bogden R., Ederle D., Stella A., Vecchietti A., Kater

M.M., Masiero S., Lasserre P., Lespinasse Y., Allan A.C., Bus V., Chagne D., Crowhurst R.N., Gleave A.P., Lavezzo E., Fawcett J.A., Proost S., Rouze P., Sterck L., Toppo S., Lazzari B., Hellens R.P., Durel C.-E., Gutin A., Bumgarner R.E., Gardiner S.E., Skolnick M., Egholm M., Van de Peer, Yves, Salamini F., Viola R. (2010) The genome of the domesticated apple (Malus x domestica Borkh.). *Nature Genetics*, 42(10), 833-839.

Verde I., Jenkins J., Dondini L., Micali S., Pagliarani G., Vendramin E., Paris R., Aramini V., Gazza L., Rossini L., Bassi D., Troggio M., Shu S., Grimwood J., Tartarini S., Dettori M.T., Schmutz J. (2017) The Peach v2.0 release: high-resolution linkage mapping and deep resequencing improve chromosome-scale assembly and contiguity. *BMC Genomics*, 18(1), 225.

Verma V., Ravindran P., Kumar P.P. (2016) Plant hormone-mediated regulation of stress responses. *BMC Plant Biology*, 16(1), 86.

Vidhyasekaran P. (2015) *Plant Hormone Signaling Systems in Plant Innate Immunity*, Vol. 2, Springer, Dordrecht, Netherlands.

Wang X., Jiang N., Liu J., Liu W., Wang G.-L. (2014) The role of effectors and host immunity in plant–necrotrophic fungal interactions. *Virulence*, 5(7), 722–732.

Wang Z., Gerstein M., Snyder M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews. Genetics*, 10(1), 57–63.

Wanjugi H., Coleman-Derr D., Huo N., Kianian S.F., Luo M.-C., Wu J., Anderson O., Gu Y.Q. (2009) Rapid development of PCR-based genome-specific repetitive DNA junction markers in wheat. *Genome*, 52(6), 576–587.

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., Birch, Paul R. J. (2007) A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature*, 450(7166), 115-118.

Whitaker V.M., Bradeen J.M., Debener T., Biber A., Hokanson S.C. (2010) Rdr3, a novel locus conferring black spot disease resistance in tetraploid rose: Genetic analysis, LRR profiling, and SCAR marker development. *Theoretical and Applied Genetics*, 120(3), 573–585.

Wildermuth M.C., Dewdney J., Wu G., Ausubel F.M. (2001) Isochorismate synthase is required to synthesize salicylic acid for plant defence. *Nature*, 414(6863), 562–565.

Williams T.A., Nakjang S., Campbell S.E., Freeman M.A., Eydal M., Moore K., Hirt R.P., Embley T.M., Williams B.A.P. (2016) A Recent Whole-Genome Duplication Divides Populations of a Globally Distributed Microsporidian. *Molecular Biology and Evolution*, 33(8), 2002–2015.

Wissemann V. (2003) Conventional Taxonomy (Wild Roses). In: A. V. Roberts, T. Debener, S. Gudin (Eds). *Encyclopedia of rose science*. Elsevier Academic press, Oxford: 111–117.

Wissemann V., Ritz C.M. (2005) The genus Rosa (Rosoideae, Rosaceae) revisited: molecular analysis of nrITS-1 and atpB-rbcL intergenic spacer (IGS) versus conventional taxonomy. *Botanical Journal of the Linnean Society*, 147(3), 275–290.

Wolf F.A. (1912) The perfect stage of Actinonema rosae. *Botanical Gazette*, 54(3), 218–234.

Wolf J. B. W. (2013) Principles of transcriptome analysis and gene expression quantification: an RNA-seq tutorial. *Molecular Ecology Resources*, 13(4), 559–572.

Xu Q., Wen X., Deng X. (2005) Isolation of TIR and nonTIR NBS–LRR resistance gene analogues and identification of molecular markers linked to a powdery mildew resistance locus in chestnut rose (Rosa roxburghii Tratt). *Theoretical and Applied Genetics*, 111(5), 819–830.

Yamamoto M., Nakatsuka S., Otani H., Kohmoto K., Nishimura S. (2000) (+)-catechin acts as an infection-inhibiting factor in strawberry leaf. *Phytopathology*, 90(6), 595–600.

Yang F., Jensen J.D., Svensson B., Jørgensen H.J.L., Collinge D.B., Finnie C. (2012) Secretomics identifies Fusarium graminearum proteins involved in the interaction with barley and wheat. *Molecular Plant Pathology*, 13(5), 445–453.

Yang Y.-X., Ahammed G.J., Wu C., Fan S.-y., Zhou Y.-H. (2015) Crosstalk among Jasmonate, Salicylate and Ethylene Signaling Pathways in Plant Disease and Immune Responses. *Current Protein & Peptide Science*, 16(5), 450–461.

Yoshida K., Saitoh H., Fujisawa S., Kanzaki H., Matsumura H., Yoshida K., Tosa Y., Chuma I., Takano Y., Win J., Kamoun S., Terauchi R. (2009) Association genetics reveals three novel avirulence genes from the rice blast fungal pathogen Magnaporthe oryzae. *Plant Cell*, 21(5), 1573–1591.

Zhang L., Du L., Poovaiah B.W. (2014) Calcium signaling and biotic defense responses in plants. *Plant Signaling & Behavior*, 9, e973818.

Zhang S.-H., Yang Q., Ma R.-C. (2007) Erwinia carotovora ssp carotovora Infection Induced “Defense Lignin” Accumulation and Lignin Biosynthetic Gene Expression in Chinese Cabbage (Brassica rapa L. ssp. pekinensis). *Journal of Integrative Plant Biology*, 49(7), 993–1002.

Zhang Y.L., Fan W.H., Kinkema M., Li X., Dong X.N. (1999) Interaction of NPR1 with basic leucine zipper protein transcription factors that bind sequences required for salicylic acid induction of the PR-1 gene. *Proceedings of the National Academy of Sciences*, *96*(11), 6523–6528.

Zheng W., Chung L.M., Zhao H. (2011) Bias detection and correction in RNA-Sequencing data. *BMC Bioinformatics*, 12(1), 290.

Zhou J.M., Trifa Y., Silva H., Pontier D., Lam E., Shah J., Klessig D.F. (2000) NPR1 differentially interacts with members of the TGA/OBF family of transcription factors that bind an element of the PR-1 gene required for induction by salicylic acid. *Molecular Plant-Microbe Interaction*, 13(2), 191–202.

Zimin A., Stevens K.A., Crepeau M.W., Holtz-Morris A., Koriabine M., Marçais G., Puiu D., Roberts M., Wegrzyn J.L., Jong P.J. de, Neale D.B., Salzberg S.L., Yorke J.A., Langley C.H. (2014) Sequencing and assembly of the 22-gb loblolly pine genome. *Genetics*, 196(3), 875–890.

Zipfel C., Felix G. (2005) Plants and animals: a different taste for microbes? *Current Opinion in Plant Biology*, 8(4), 353–360.

Zlesak D.C. (2006) ROSE: Rosa x hybrida. In: Anderson N. O. (Ed). *Flower breeding and genetics: Issues, challenges and opportunities for the 21st century*. Springer, Dordrecht, Netherlands: 695–738.

Appendices:

Supplementary Table 1: Overview of the inoculation experiments performed for generating the MACE- and RNAseq-data

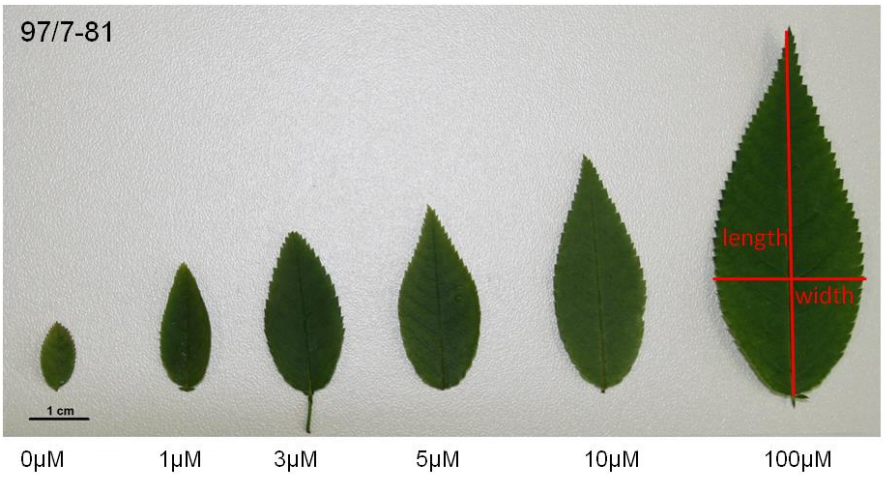
Sample name	Treatment	BR1	BR2	BR3	BR4	Sequencing approach	Chapter
Start	PC leaf extracted from the plant and directly frozen in liquid nitrogen	0 hpi	0 hpi	0 hpi	-	MACE	6
PC-control	PC leaf extracted from the plant and stored in a translucent plastic boxes, comparable to inoculated samples	0 hpi	0 hpi	0 hpi	-	MACE	4
		24 hpi	24 hpi	24 hpi	-	MACE	4
		72 hpi	72 hpi	72 hpi	-	MACE	4
PR+DR	PC leaf inoculated with spore solution of the *D. rosae* isolate DortE4 (compatible interaction)	0 hpi	0 hpi	0 hpi	-	MACE	4
		24 hpi	24 hpi	24 hpi	-	MACE	4
		72 hpi	72 hpi	72 hpi	-	MACE/ RNAseq	4/ 6
PC+PP	PC leaf inoculated with a multispore isolate (greenhouse mixture) of *P. pannosa* (compatible interaction)	0 hpi	0 hpi	0 hpi	-	MACE	4
		24 hpi	24 hpi	24 hpi	-	MACE	4
		72 hpi	72 hpi	72 hpi	-	MACE	4
91+DR	91/100-5 leaf inoculated with spore solution of the *D. rosae* isolate DortE4 (incompatible interaction)	-	24 hpi	24 hpi	24 hpi	MACE	6
		-	72 hpi	72 hpi	72 hpi	MACE/ RNAseq	6
RGA1+DR	muRdr1A leaf inoculated with spore solution of the *D. rosae* isolate DortE4 (incompatible interaction)	-	24 hpi	-	-	MACE	6
		-	72 hpi	-	-	MACE	6

Supplementary Table 2: Fertiliser composition. The solution was prepared as a 10 or 20 fold stock solution and diluted for the application.

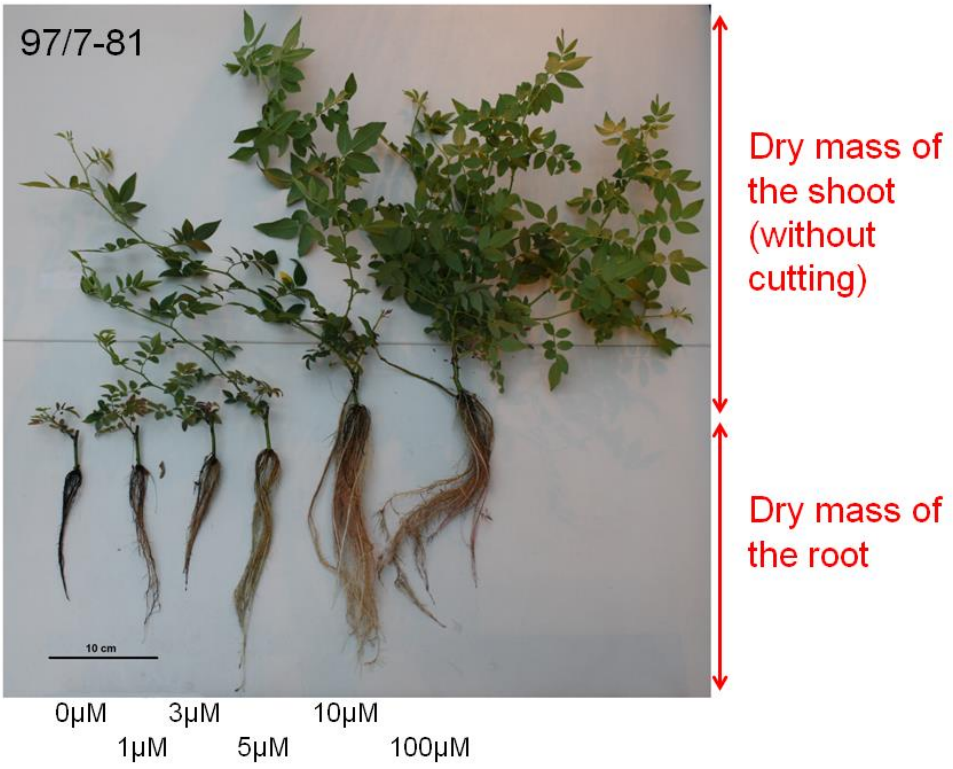
Component	Mass fraction [g/l]	Concentration [mmol/l]
KNO_3	0.22	2.176
NH_4NO_3	0.255	3.186
K_2SO_4	0.075	0.430
$MgSO_4$*7H_2O	0.15	0.609
KH_2PO_4	0.00136	0.001
Ferty 10® trace elements fertiliser (Planta Düngemittel, Regenstauf, Germany)	0.1	-

Supplementary Table 3: Composition of the nutrient solution of the starvation experiments. The solution was prepared as a 25 fold stock solution and diluted for the application.

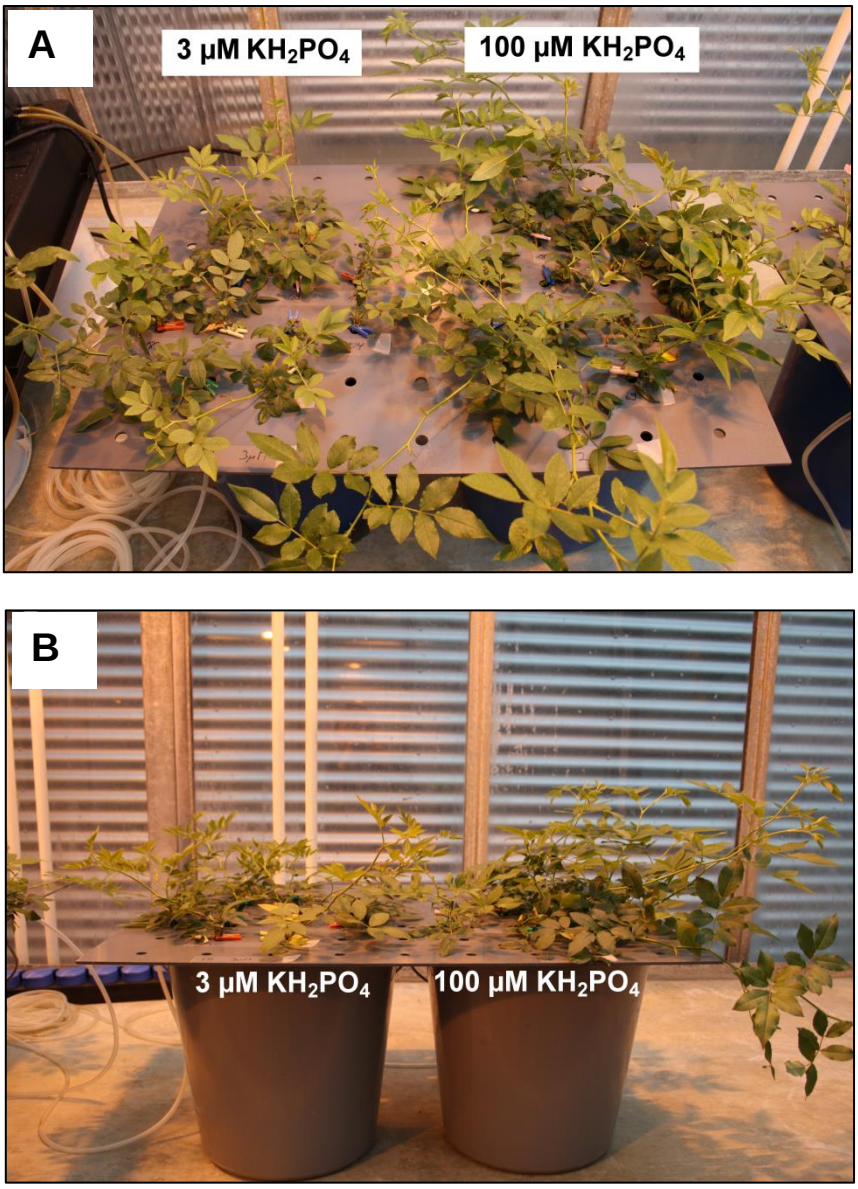
Component	Mass fraction [mg/l]	Concentration [mmol/l]
NH_4NO_3	24.0120	0.3
$Ca(NO_3)_2$+4H_2O	318.8025	1.35
$CaCl_2$+2H_2O	22.0530	0.15
K_2SO_4	130.695	0.75
$MgSO_4$+7H_2O	24.6480	0.1
$MgCl_2$+6H_2O	132.145	0.65
$ZnSO_4$+7H_20	0.2875	0.001
$MnSO_4$+H_2O	0.3380	0.002
$CuSO_4$+5H_2O	0.0025	0.0001
H_3BO_3	0.1237	0.002
$NaMoO_4$+2H2O	0.1210	0.0005
Fe-EDTA (Fetrilon® 13%, Compo, Münster ,Germany)	11.0115	0.03
2-(N-morpholino)ethanesulfonic acid (MES)	149.2505	0.7
KH_2PO_4	13.609	0.1
KH_2PO_4	1.3609	0.01
KH_2PO_4	0.6805	0.005
KH_2PO_4	0.4083	0.003
KH_2PO_4	0.1361	0.001



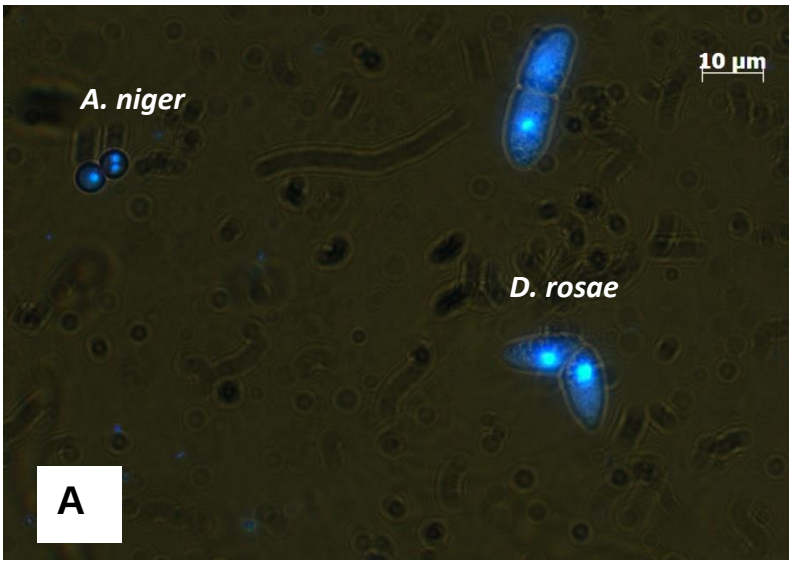
Supplementary Figure 1: Leaf parameter as indicators of P-starvation.

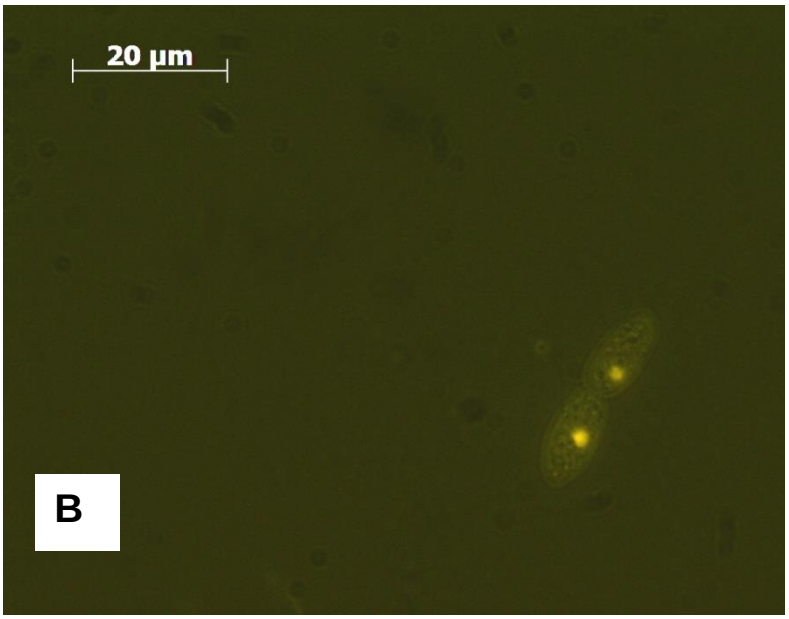


Supplementary Figure 2: Calculation of the root/shoot ration based on the dry mass as indicators of P-starvation.



Supplementary Figure 3: 97/7 genotypes grown for 5 weeks under 3 µM and 100 µM P-treatment.





Supplementary Figure 4: Stained nuclei of *D. rosae* within spores. (A) DAPI staining of spores of *D. rosae* and *Aspergillus niger* as a standard (100x), (B) PI staining of *D. rosae* spores (100x)

Electronical appendix

Electronic supplementary material ES1: Assembly and annotation of the 88 genome. (DVD 1)

Electronic supplementary material ES2: BUSCO results of the genome sequence of 88 genome. (DVD 2)

Electronic supplementary material ES3: Results of the MACE and RNAseq data. (DVD 2)

Electronic supplementary material ES4: Potential contigs of *P. pannosa*. This file contains the assembly as well as the annotations of the contigs potentially originated from the fungus *P. pannosa*. Additionally, the results of the Full_lengtherNEXT package and BLAST to sequences of *P. xantii* are included. (DVD 2)

Electronic supplementary material ES5: Data of indicators for P-starvation symptoms (DVD 2)

Electronic supplementary material ES6: Supplemental material of manuscript I (DVD 2)

Electronic supplementary material ES7: Supplemental material of manuscript II (DVD 2)

Electronic supplementary material ES8: Supplemental material of manuscript III (DVD 2)

Curriculum Vitae

Personal information

Name:	Enzo Neu
Birth name:	Klein
Date of birth:	20.01.1989
Place of birth:	Saarbrücken (Germany)
Nationality:	German

Employment History

05.2018-today	Research assistant at KWS SE

Academic education

10.2013-08.2018	PhD student at the Leibniz Universität Hannover, Institute of Plant Genetics, Department of Molecular Plant Breeding (Prof. Thomas Debener) Research Training Group GRK1798 “Signaling at the Plant-Soil Interface” Thesis: Genomic and transcriptomic analysis of the interaction of roses with the black spot fungus *Diplocarpon rosae*
10.2011-09.2013	Student at the Leibniz Universität Hannover M.Sc. Plant Biotechnology Thesis: Analysis of potential effector candidates in the genome of *Diplocarpon rosae*.
10.2008-09.2013	Student at the Leibniz Universität Hannover B.Sc. Plant Biotechnology Thesis: Mapping of microsatellite markers in a segregating rose population

School education

08.2001-06.2008	Secondary School: Albert-Schweitzer-Gymnasiums Dillingen

	Allgemeine Hochschulreife
08.1999-06.2001	Secondary School: Max-Planck-Gymnasiums Saarlouis
08.1995-07.1999	Primary School: Don Bosco Grundschule Neuforweiler

Publications

Neu, E.; Domes, H.; Kaufmann, H.; Linde, M.; Menz, I.; Debener, T. (2018): Interaction of roses with a biotrophic and a hemibiotrophic leaf pathogen leads to differences in defense transcriptome activation. Plant Molecular Biology, submitted

Neu, E.; Debener, T. (2018: Prediction of the *Diplocarpon rosae* secretome reveals candidate genes for effectors and virulence factors. Fungal Biology, submitted

Hibrand Saint-Oyant, L.; Ruttink, T.; Hamama, L.; Kirov, I.; Lakwani, D.; Zhou, N.-N.; Bourke, P.M.; Daccord, N.; Leus, L.; Schulz, D.; Van de Geest, H.; Hesselink, T. ; Van Laere K.; Debray; K. ; Balzergue S.; Thouroude T.; Chastellier A.; Jeauffre J.; Voisine, L.; Gaillard, S.; Borm, T.J.A.; Arens, P.; Voorrips, R.E.; Maliepaard, C.; Neu, E.; Linde, M.; Le Paslier, M.C.; Bérard, A.; Bounon, R. ; Clotault, J.; Choisne, N.; Quesneville, H.; Kawamura, K.; Aubourg, S.; Sakr, S.; Smulders, M.J.M.; Schijlen, E.; Bucher, E.; Debener, T.; De Riek, J.; Foucher, F. (2018): A high-quality sequence of *Rosa chinensis* to elucidate genome structure and ornamental traits. Nature Plants (4), 473-484.

Neu, E. Rees, J.; Featherston, J.; Debener, T. (2017): A draft genome sequence of the rose black spot fungus Diplocarpon rosae reveals a high degree of genome duplication. PLoS ONE 12(10) (Artikel)

Conference Contributions

Oral presentation

Klein, E.; Kaufmann, H.; Debener, T. (2016): Transcriptome Analysis of the Defense Response of Roses in the Interaction with *Diplocarpon rosae* and *Podosphaera pannosa* using the MACE Technique. PAG XXVI – Plant & Animal Genome Conference, San Diego (USA)

Klein, E.; Kaufmann, H.; Debener, T. (2015): Transkriptomanalyse der Interaktion von Rosen mit *Diplocarpon rosae* und *Podosphaera pannosa* mit Hilfe der MACE-Technik. Tagung der GPZ-AG 18 (Zierpflanzen), Hannover (Germany)

Posters

Klein, E.; Kaufmann, H.; Domes, H.; Debener, T. (2017): Transcriptome analysis of the defense response of roses in the interaction with *Diplocarpon rosae* and

Podosphaera pannosa using the MACE technique. Seventh international symposium on rose research and cultivation, Angers (France)

Klein, E.; Rees, J.; Featherston, J.; Debener, T. (2016): Effector-Mining in the genome of Diplocapon rosae. RGC 8 - 8th International Rosaceae Genomics Conference, Angers (France)

Klein, E.; Kaufmann, H.; Debener, T. (2015): RNA-seq Analyse der Interaktion von Rosen mit *Diplocarpon rosae* und *Podosphaera pannosa* mit Hilfe der MACE-Technik. Workshop der Gesellschaft für Pflanzenbiotechnologie, Einbeck (Germany)

Klein, E.; Kaufmann, H.; Linde, M.; Debener, T. (2014): RNAseq-based analysis of defence-related genes in the interaction of roses with *Diplocarpon rosae* and *Podosphaera pannosa*. GPZ-Haupttagung, Kiel (Germany)

Klein, E.; Rees, J.; Debener, T.: Analysis of potential effector genes in the genome sequence of *Diplocarpon rosae*. Sixth international symposium on rose research and cultivation, Hannover (Germany)

Danksagung

Ich möchte diese Gelegenheit nutzen, um mich bei allen Menschen zu bedanken, die mich während meiner Promotion begleitet und unterstützt haben.

Lieber Thomas, dir gebührt mein besonderer Dank, da du mich während der Bearbeitung des Themas immer mit großem Engagement und Einsatz gefördert hast und mir das Vertrauen geschenkt hast, meine eigenen Wege zu gehen. Deine Begleitung hat mir viele Türen geöffnet und mir geholfen, mich auf vielen Ebenen weiterzuentwickeln. Ich danke dir, dass du immer ein offenes Ohr hattest, auch abseits der fachlichen Diskussionen.

Ein weiterer großer Dank gebührt allen meinen Kollegen in der Abteilung für Molekulare Pflanzenzüchtung. Dank euch hat mir die Arbeit stets sehr viel Spaß gemacht. Ihr habt mir mit Rat und Tat zur Seite gestanden und dafür gesorgt, dass ich mich stets sehr wohl gefühlt habe.

Besonders hervorheben möchte ich hier meine Mitdoktoranden Ina Menz, Annette Bartkiewicz, Friederike Chilla, Helena Domes, Juliane Geike und Vinicius Vilperte. Ihr wart einer der Hauptgründe, dass diese Zeit so außergewöhnlich war. Mit euch habe ich viel gelacht und gescherzt. Wir haben uns immer miteinander gefreut und uns gegenseitig motiviert. Gleichzeitig haben wir uns immer unter- und auch gestützt, wenn das nötig war. Dafür mein tiefes Dankeschön.

Des Weiteren möchte ich mich bei Prof. Helge Küster für das Koreferat und Dr. Sascha Offermann für den Prüfungsvorsitz bedanken, sowie bei allen anderen Mitgliedern des GRK 1798 für die schönen Erfahrungen, die ich innerhalb dieses Netzwerkes machen konnte. Ein weiterer Dank geht an Jonathan Featherston und Jasper Rees für die gute Zusammenarbeit an der Genomsequenz von *D. rosae* und Fabrice Foucher für die Möglichkeit, Teil des Sequenzierprojekts der Rose zu sein.

Zum Abschluss möchte ich mich bei meinen Freunden und meiner Familie bedanken, die mich immer unterstützt und motiviert haben. Vor allem ohne die unermüdliche Unterstützung meiner Mutter wäre ich nicht an dem Punkt, an dem ich heute bin. Und zuletzt gebührt mein Dank meiner Frau Jana Neu, die immer an mich geglaubt und hinter mir gestanden hat und mir geholfen hat, den Weg zu finden, der für uns der Richtige ist.