PST SECRETOME PROJECT, A SHORT CUT TO EFFECTOR ANALYSIS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

AHMET ÇAĞLAR ÖZKETEN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
BIOTECHNOLOGY

JANUARY 2019

Approval of the thesis:

PST SECRETOME PROJECT, A SHORT CUT TO EFFECTOR ANALYSIS

submitted by **AHMET ÇAĞLAR ÖZKETEN** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Biotechnology Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** _____

Assoc. Prof. Dr. Can Özen
Head of Department, **Biotechnology** _____

Prof. Dr. Mahinur S. Akkaya
Supervisor, **Biotechnology, METU** _____

Assoc. Prof. Dr. Bala Gür Dedeoğlu
Co-Supervisor, **Biotechnology, Ankara University** _____

Examining Committee Members:

Assoc. Prof. Dr. Özlen Konu
Molecular Biology and Genetics, Bilkent University _____

Prof. Dr. Mahinur S. Akkaya
Biotechnology, METU _____

Assoc. Prof. Dr. Çağdaş D. Son
Biology, Middle East Technical University _____

Assoc. Prof. Dr. Aslıhan Günel
Chemistry, Ahi Evran University _____

Assoc. Prof. Dr. Mehmet Somel
Biology, Middle East Technical University _____

Date: 15.01.2019

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Ahmet Çağlar Özketen

Signature:

ABSTRACT

PST SECRETOME PROJECT, A SHORT CUT TO EFFECTOR ANALYSIS

Özketen, Ahmet Çağlar
Doctor of Philosophy, Biotechnology
Supervisor: Prof. Dr. Mahinur S. Akkaya
Co-Supervisor: Assoc. Prof. Dr. Bala Gür Dedeoğlu

January 2019, 167 pages

Wheat yellow (stripe) rust is a devastating disease on wheat on the global scale. Emergence of new aggressive races leads swift epidemics unless countermeasures are in place. The causative agent of the disease is an obligate biotrophic fungus, *Puccinia striiformis* f.sp. *tritici* (*Pst*). The identification of the infection mechanisms is the best strategy to act against the disease. For that reason, candidate secreted effector proteins (CSEPs) of Pst races in Turkey was monitored using *de novo* transcriptome sequencing approach combined with data mining on the inoculated tissues of compatible and incompatible interaction. The small-secreted proteins (SSPs) of the identified differentially expressed genes (DEGs) were generated and characterized *via* in silico analysis. Comparisons with published reports reveal both unique and common sets of candidate effector for virulence. In order to validate reliability of the data, some of the effector candidates were cloned and studied for their function. Unigene17495 (Pstg10917) targets chloroplasts and suppresses the cell death triggered by INF1 elicitor in planta studies. The thesis includes the microarray profiling of differentially expressed genes during compatible and incompatible interaction. Moreover, the new screening methods for determination of biological functions were studied in the thesis in order to pinpoint flexible strategies for screening effectors.

Keywords: Secretome, transcriptome, effectors, Puccinia striiformis, virulence, pathogenicity, yellow/stripe rust, microarray

ÖZ

PST SEKRETOM PROJESİ, EFEKTÖR ÇALIŞMALARINA KISA YOL

Özketen, Ahmet Çağlar
Doktora, Biyoteknoloji
Tez Danışmanı: Prof. Dr. Mahinur S. Akkaya
Ortak Tez Danışmanı: Doç. Dr. Bala Gür Dedeoğlu

Ocak 2019, 167 sayfa

Dünya çapında, sarı pas hastalığı buğday için yıkıcı etkidedir. Eğer gerekli önlemler alınmazsa, yeni patojen ırklarının ortaya çıkması hızla yayılarak epidemiklere yol açabilir. Hastaığı sebebi obligat biyotrofik bir mantar olan *Puccinia striiformis* f.sp. *tritici* (*Pst*) patojenidir. Enfeksiyon mekanizmasının nasıl gerçekletiğinin tespiti, hastalığa karşı eldeki en iyi stratejidir. Bu nedenle aday effektör proteinleri, dirençli ve dirençsiz bitkilerin patojen inoküle edilmiş dokularını kullanarak data madenciliği ile harmanlanmış de novo transkriptom dizilemesi yolu ile tespit edildi. Karşılaştımalı anlatım düzeyi analizi ile tespit edilmiş genlerin içerisinde küçük sekretom proteinleri biyoinformatik analizlerle saptandı. Literatürdeki raporlarla yapılan karşılaştırmalar yeni ya da ortak virülansta sorumlu efektör adaylarının bulunmasını sağladı. Verinin güvenilirliği test etmek için bazı aday efektörler klonlandı ve fonksiyon açısından karakterize edildi. Çalışmalar sonucunda Unigene17495 (Pstg10917) isimli aday efektörün bitki hücresinde kloroplastı hedeflediği ve INF1 proteini sebepli hücre ölümlerini ise baskıladığı gözlemlendi. Bu tez kapsamında dirençli ve dirençsiz bitkilerdeki inokülasyon sonrası değişen patojen gen anlatım düzeyleri mikro dizin analizi ile belirlendi. Aynı şekilde, tespit edilen adayların en uygun şekilde fonksiyon analizlerinin yapılabilmesi yeni nesil metodolojiler üzerinde de çalışılma yapılmıştır.

Anahtar Kelimeler: Sekretom, transkriptom, efektör, Puccinia striiformis, virülans, patojenisite, sarı pas, mikro dizin

Dedicated to my father and my wife

ACKNOWLEDGMENTS

I would like to express my grandest gratitude to my supervisor Prof. Dr. Mahinur S. Akkaya. She guides me to find my best potential as I cruise through my own scientific journey. Her endless energy motivates me to reach the sky in terms of creativity and resourcefulness.

I am thankful to Assoc. Prof. Dr. Bala Gür Dedeoğlu for backing me up as my co-supervisor whenever I need. Furthermore, she and her student, Dr. Yasemin Öztemur, assist me during the microarray analysis. I also owe an earnest gratitude to my thesis monitoring comitee members, Dr. Özlen Konu and Dr. Çağdaş D. Son as they directed me to enrich my thesis work. I am sincerely appreciative to my remaining thesis defense comitee members Dr. Aslıhan Günel and Dr. Mehmet Somel for their encouragement.

My thesis work is under the subject of prioritized areas in science declared by TUBITAK. Hence, I am grateful to TUBITAK BIDEB 2211/C thesis scholarship for the support. I am thankful for the grants TUBITAK 110T445, TUBITAK 113Z350 and METU BAP.

I am grateful for the bacteria and the vectors obtained as gifts from Dr. Tolga O. Bozkurt, Dr. Sophien Kamoun, Dr. Joe Win, Dr. Narayana Upadhyaya and Dr. Yasin Dağdaş.

I also want to state my gratitudes to my labmates, my friends and family members, especially to Dr. Bayantes Dagvadorj, Zemran Mustafa, Murat Erdem, Burak Demiralay, Dorukcan Özköse and the others that I couldn't write due to limited space issues.

Last and foremost, I am appreciative for all of the hard work and supports of my beautiful wife, lab. partner and soulmate Dr. Ayşe Andaç during my PhD thesis and life troubles. Without whom, I couldn't achieve this success.

Our words are 'Scientia Dux Vitae Certissimus'.

TABLE OF CONTENTS

ABSTRACT v

ÖZ vii

ACKNOWLEDGMENTS x

TABLE OF CONTENTS xi

LIST OF TABLES xv

LIST OF FIGURES xvii

LIST OF ABBREVIATIONS xix

LIST OF SYMBOLS xxiii

CHAPTERS

1.1 Wheat 1

1.2 Wheat stripe (yellow) rust 3

1.3 Plant immunity 6

1.4 Effectors 12

1.4.1 *Pst* effectors 17

1.5 Strategies for the discovery and the investigation of the effectors 20

1.5.1 Genomics, transcriptomics and proteomics 20

1.5.2 Data mining 25

1.5.3 Type 3 secretion system based effector screening 31

1.5.4 Host induced gene silencing 32

1.5.5 Heterologous Systems via *Agrobacterium tumefaciens* Mediated Gene Transfer 33

1.5.6 Subcellular localization studies 35

1.6 Aim of the research 36
2.1. Plant and pathogen materials, growth conditions 39
2.2 Tissue sampling and RNA isolation for sequencing 40
2.3 Transcriptome sequencing 41
2.4 *De novo* transcriptome assembly 43
2.5 Transcriptome annotation 43
2.6 Analysis of expression difference on transcriptome data 45
2.7 Mapping of *Pst* related genes and novel genes 46
2.8 Construction of differentially expressed *Pst*-secretome 48
2.9 *In silico* characterization of PstDESSPs 48
2.9.1 Function annotations 49
2.9.2 Subcellular localization predictions 49
2.9.3 Conserved domain analysis 50
2.9.4 Virulence analysis 50
2.9.5 Effector predictions 51
2.9.6 Comparisons with published works 51
2.9.7 Phylogenetic tree analysis 52
2.9.8 Motif discoveries 53
2.10 RNA isolation for microarray assays 53
2.11 Microarray analysis 54
2.12 Pilot studies for candidate effector characterizations *in vivo* 55
2.12.1 *P. fluorescens* EtHAn mediated wheat infiltration assays 55
2.12.2 Optimizations of WIA assays 59
2.12.3 Cloning of eGFP into pEDV6 vector 60

2.12.4 *P. fluorescens* EtHAn mediated PTI assays....66
2.13 Amplification and cloning of candidate effectors....68
2.13.1 cDNA synthesis....69
2.13.2 Amplification of candidate effectors from cDNA templates....70
2.13.3 Cloning to pK7FWG2 vector with Gateway cloning....71
2.13.4 Transformation to *A. tumefaciens*....72
2.14 *A. tumefaciens* mediated cell death assays....73
2.15 Subcellular localization experiments....74
3.1 *P. fluorescens* EtHAn mediated wheat infiltration assays....78
3.2 *P. fluorescens* EtHAn mediated suppression assays....83
3.3 Transcriptome sequencing....87
3.4 De novo assembly....87
3.5 Annotation and classification....88
3.6 Identification and mapping of DEGs....89
3.7 Differentially expressed small secreted proteins of *Pst* (PstDESSPs)....93
3.8 *In silico* characterization of PstDESSPs....95
3.9 Findings of microarray analysis....99
3.10 Cell death suppression assays of candidate effectors....101
3.11 Subcellular localization analysis of candidate effectors....104
REFERENCES....113
APPENDICES
A. Plasmid Maps....133
B. List of The Organelle Targetting Proteins....135
C. Annotation of PstDESSPs....136

D. Characterization of PstDESSPs 144
E. Effector Prediction Among PstDESSPs 151
F. Comparisons of PstDESSPs with Other Reports 159
G. Length Distribution of Contigs and Unigenes 166
CURRICULUM VITAE 167

LIST OF TABLES

TABLES

Table 1.1. Global wheat production and agriculture areas comparisons. (FAOstat 2016 data).....2
Table 1.2. The list of Pst effectors reported in literature.....18
Table 1.3. List of the 'Omics' studies conducted on Pst.....23
Table 1.4. Valuable databases and prediction programs used in effector mining.....28
Table 2.1. The databases used to annotate Unigenes. 44
Table 2.2. Ingredients for LB and LB agar preparations.....56
Table 2.3. Recipe for minimal media preparation.....57
Table 2.4. YR differential lines 58
Table 2.5. YR differential lines of Avocet near isogenic lines. 59
Table 2.6. Polymerase Chain Reaction (PCR) used for amplification of eGFP. 60
Table 2.7. Reaction Parameters for Thermal Cycler. 61
Table 2.8. The components of pENTR/D-TOPO cloning reaction. 61
Table 2.9. List of the reactants used in colony PCR. 63
Table 2.10. Parameters for the setup of Thermocycler machine. 63
Table 2.11. The reagents for LR cloning. 65
Table 2.12. Primers used in this study.....69
Table 2.13. The ingredients for DNase I treatment. 69
Table 2.14. The list of the reagents used for first strand cDNA synthesis. 70
Table 3.1. The cell death counts observed on wheat differential lines after WIA. ...80
Table 3.2. The replication of WIA assay on Avocet near isogenic lines. 81
Table 3.3. The results for the final trial of WIA using the effector candidates on wheat cultivars. 83
Table 3.4. Statistics of collected clean reads of compatible and incompatible interaction.....87
Table 3.5. De novo assembly statistics for Pst inoculated/infected wheat transcriptome. 88

Table 3.6. Annotation statistics of the assembled unigenes of transcriptome sequencing project. 88

LIST OF FIGURES

FIGURES

Figure 1.1. Wheat production of Turkey compared to leading 10 wheat producers...2
Figure 1.2. The Life Cycle of Puccinia striiformis f.sp. tritici (Chen 2017).....4
Figure 1.3. Schematic overview of plant innate immunity. 8
Figure 1.4. A graphical abstract to summarize current models used to elucidate Avr-R protein interactions in plant innate immunity.....11
Figure 1.5. Production, delivery and trafficking of effectors via haustoria (H) interface of Pst upon infection. 16
Figure 2.1. Pst inoculation on two leaf stage wheat leaves.....39
Figure 2.2. The flowchart of the design of the experiments conducted for tissue sampling. 40
Figure 2.3. Flowchart showing the preparation of total RNA samples for transcriptome sequencing.....42
Figure 2.4. Flowchart describes the strategy to identify PstDEGs.....47
Figure 3.1. Cell death observed on wheat cultivar, Kalyansona (YR2).....78
Figure 3.2. Closer look on the cell death region observed on wheat cultivar, Kalyansona (YR2).....79
Figure 3.3. The photos of the final WIA assay trial on wheat cultivar, Kalyansona (YR2). 82
Figure 3.4. P. fluorescens T3SS delivered PstHa12J12 area is challenged with P. syringae DC3000 on N. benthamiana. 84
Figure 3.5. P. fluorescens T3SS delivered PstHa15N21 and PstHa12H2 area were challenged with P. syringae DC3000 on N. benthamiana.....84
Figure 3.6. P. fluorescens T3SS delivered candidates and P. syringae DC3000 on wheat cultivar, Avocet S (72 hpi). 86
Figure 3.7. Statistics of NR classification of unigenes. 89
Figure 3.8. Distribution of differentially expressed unigenes (DEGs); 6851 up, 3699 down regulated. 90

Figure 3.9. Functional annotations of PstDEGs via Blast2Go program. 93
Figure 3.10. The graphical comparison for the predicted small-secreted proteins (SSPs) and small-secreted candidate effectors (SSCEs) of PstDEGs and other Pucciniale proteome..... 94
Figure 3.11. Annotated PstDESSPs through similarity search against databases. ... 96
Figure 3.12. The phylogenetic tree constructed among PstDESSPs using sequence similarities..... 98
Figure 3.13. The 10 highest upregulated Pst genes generated at 24 hpi of the disease progress. 100
Figure 3.14. The 10 highest upregulated Pst genes generated at 10 dpi of the disease progress. 100
Figure 3.15. Inf1 triggered 'cell death suppression' assay. 102
Figure 3.16. P. syringae DC3000 (A) and PstSCR1 (B) triggered 'cell death suppression' assay. 103
Figure 3.17. Subcellular localization of Pstg10917-GFP (Unigene17495) inside N. benthamiana leaves. 106
Figure 3.18. Subcellular localization of Pstg11651-GFP inside N. benthamiana leaves. 107

LIST OF ABBREVIATIONS

ABBREVIATIONS

Avr:	Avirulence
BGI:	Beijing genome institute
BIC:	Biotrophic interfacial complex
BLAST:	Basic local alignment search tool
CaMV:	Cauliflower mosaic virus
CAZymes:	Carbohydrate active enzymes
CBM14:	Chitin binding motif 14
CC:	Coiled coil
CDD:	Conserved domains database
cDNA:	Complementary deoxyribonucleic acid
CEBiP:	LysM-containing chitin oligosaccharide elicitor binding protein
CERK1:	LysM-containing chitin elicitor receptor kinase 1
CRN:	Crinkler
CSEP:	Candidate secreted effector protein
CWDE:	Cell wall degrading enzymes
DAMP:	Danger associated molecular patterns
DEG:	Differentially expressed genes
Dpi:	Days post inoculation
dsRNA:	Double stranded ribonucleic acid
Ecp6:	Extracellular protein 6

EST:	Expressed sequence tags
EtHAn:	Effector to host analyzer
ETI:	Effector triggered immunity
FAO:	Food and agriculture organization
FDR:	False discovery rate
FLS2:	Flagellin sensing 2
FPKM:	Fragments per kilobase per million reads
GFP:	Green fluorescent protein
GO:	Gene ontology
HIGS:	Host induced gene silencing
HMC:	Haustorial mother cell
Hpi:	Hours post inoculation
HR:	Hypersensitive response
HSEP:	Haustoria specific effector protein
ISEP:	Infection specific effector protein
KB:	King's broth
LB:	Lysogeny broth
LED:	Lipase engineering database
LRR:	Leucine rich region
LysM:	Lysine motif
MAMP:	Microbe associated molecular patterns
MAPK1:	Mitogen activated protein kinase 1

Mlp:	*Melampsora larici* populina
NB:	Nucleotide binding
NCBI:	The national center for biotechnology information
NEP1:	Necrosis and ethylene inducing protein 1
NGS:	Next generation sequencing
NLP:	NEP1 like protein
NLR:	NB-LRR resistance proteins
NPR1:	Non-expresser of PR genes 1
ORF:	Open reading frame
PAMP:	Pathogen associated molecular patterns
PCD:	Programmed cell death
PCR:	Polymerase chain reaction
pEDV:	Effector detector vector
Pgt:	*Puccinia graminis*
PHI:	Pathogen host interaction
PI3P:	Phospholipid phosphatidylinositol-3-phosphate
PNPi:	*Puccinia* NPR1 interactor
PpEC:	*P. pachyrhizi* effector candidate
PR:	Pathogenesis related
PRR:	Pattern recognition receptor
Pst:	*Puccinia striiformis* f.sp. *tritici*
PstDESSPs:	*Pst* Differentially expressed small secreted proteins

PTI: PAMP triggered immunity

Ptt/Pt: *Puccinia triticina*

PVPP: Polyvinylpolypyrrolidone

R: Resistance

RIN: RNA integrity number

ROS: Reactive oxygen species

SCEP: Small cysteine rich effector

Shr: Suppressor of HR

siRNA: Small interfering ribonucleic acid

SP: Signal peptide

SSP: Small secreted protein

SSSP: Species-specific small-secreted protein

T3SS: Type 3 Secretion System

TIR: Toll/Interleukin-1 receptor

TP: Transit peptide

UK: United Kingdom

USA: United States of America

USDA: US Department of Agriculture

VIGS: Virus induced gene silencing

WGS: Whole genome sequencing

WIA: Wheat infiltration assay

YR: Yellow rust

LIST OF SYMBOLS

SYMBOLS

Δ: Omitted/Deleted gene sequence

μ: Micro

#: Number

CHAPTER 1

INTRODUCTION

1.1 Wheat

Wheat is a cereal grain that consists of several cereal grasses belong to the genus *Triticum* and family *Poaceae* and their edible grains. Bread wheat (*Triticum aestivum*) and pasta wheat (*Triticum durum*) are the most important among wheat varieties in terms of food and industry. Wheat is cultivated in every continent except Antarctica and it is number one crop in terms of dedicated land area, worldwide (Table 1.1). Moreover, it is one of the 'big three crops' with regard to production volumes along with maize and rice (Shewry, 2009). In 2016, world production is over 740 million tones according to official website of FAO statistics (http://www.fao.org/faostat/en/). Turkey is placed 11th with respect to production volume in the world in 2016. According to US Department of Agriculture (USDA) estimations, Turkey will be among top 10 producers in 2018 (Figure 1.1). Thus, incontrovertible impact of wheat in our economy makes any threat against its well-being a major concern for plant biotechnology.

Table 1.1. Global wheat production and agriculture areas comparisons. (FAOstat 2016 data)

	Crops	Production (tones)	Crops	Land Area (ha)
1	Sugar cane	1890661751	Wheat	220107551
2	Maize	1060107470	Maize	187959116
3	Wheat	749460077	Rice paddy	159807722
4	Rice paddy	740961445	Soybeans	121532432
5	Potatoes	376826967	Barley	46923218
6	Soybeans	334894085	Sorghum	44771056
7	Oil palm fruit	300252193	Rapeseed	33708547
8	Sugar beet	277230790	Millet	31705489
9	Cassava	277102564	Seed cotton	30206843
10	Tomatoes	177042359	Beans dry	29392817

	Country	Production (tones)
1	China	131696392
2	India	93500000
3	Russia	73294568
4	USA	62859050
5	Canada	30486700
6	France	29504454
7	Ukraine	26098830
8	Pakistan	26005213
9	Germany	24463800
10	Australia	22274514
11	Turkey	20600000

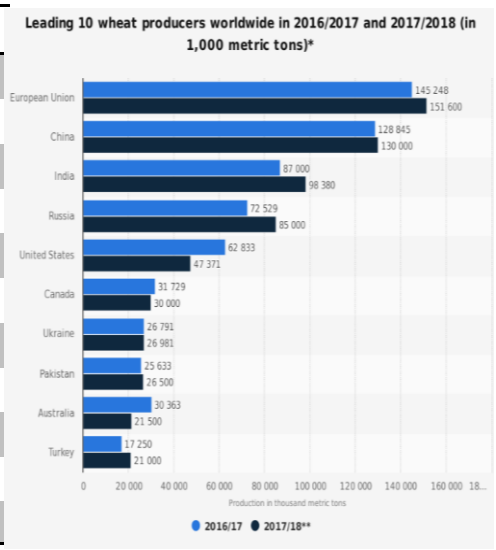


Figure 1.1. Wheat production of Turkey compared to leading 10 wheat producers. Table at the left shows 2016 production based on FAO data; whereas on the right site, leading 10 wheat producers worldwide in 2016/2017 and 2017/2018 estimations (in 1,000 metric tons) * based on US Department of Agriculture (n.d.) data. In Statista - The Statistics Portal. (Retrieved June 14, 2018, from https://www.statista.com/statistics/237908/global-top-wheat-producing-countries/.)

1.2 Wheat stripe (yellow) rust

Wheat stripe (yellow) rust is one of the most damaging diseases leading to substantial yield losses on wheat production, globally. The causative agent of the disease, *Puccinia striiformis* f. sp. (latin: 'formae speciales') *tritici* (*Pst*), is an obligate biotrophic fungus specialized on infecting wheat. *Pst* has a macrocyclic heteroecious sexual life and form five different spore stages: pycniospore (spermatium) (0), aeciospore (I), urediniospore (II), teliospore (III) and basidiospore (IV). Asexual stages (Uredinial and telial stages) of life cycle occur on primary (wheat) and secondary (other grass spp.) hosts whereas sexual life cycle occurs on alternate host (barberry spp.) (Figure 1.2)

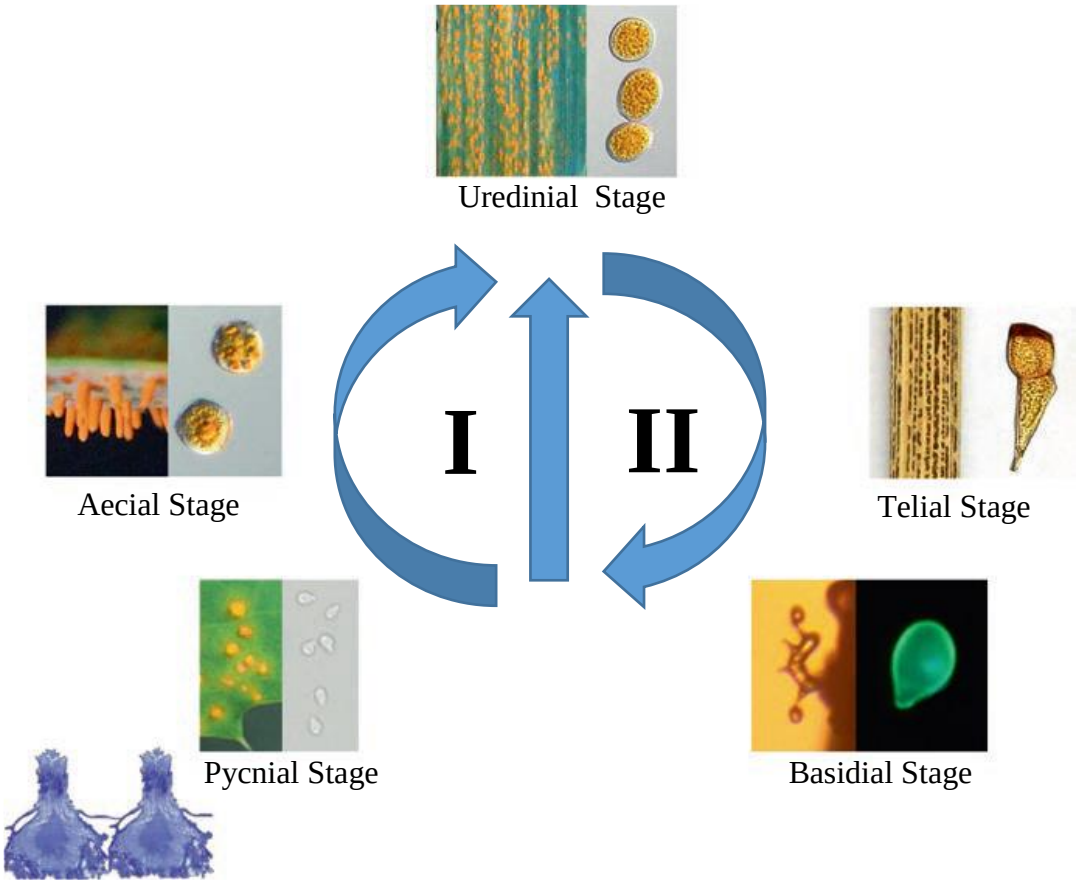


Figure 1.2. The Life Cycle of Puccinia striiformis f.sp. tritici (Chen 2017). I) Sexual cycle on alternate host and II) Asexual cycle on primary host.

Similar to the life cycle of other cereal rust fungi, *Pst* also requires two botanically distinct plant species to achieve sexual life cycle (Chen & Kang, 2017). However, the alternate host for *Pst* sexual life cycle was discovered quite recently in contrast to all efforts. Naturally formed aeciospores of *Puccinia striiformis* f. sp. *poae* obtained from *Berberis* spp. (*Berberis chinensis* and *B. koreana*) was used to infect grass species, *Poa pratensis*. Pycnia and aecia of *Pst* were produced on *Berberis* spp. under controlled conditions. Moreover, the aeciospores from B. chinensis were able to form uredinia on wheat (Jin et al., 2010). *Pst* and *P. striiformis* f.sp. *hordei* on *Berberis vulgaris* (Wang et al., 2015) and *Pst* on *Mahonia aquifolium* (taxonomically related to *Berberis*) (M. N. Wang & Chen, 2013) were able to infect under controlled

conditions as it was stated in recent reports. Sexual life cycle on barberry may help *Pst* to obtain new varieties. On the other hand, there is no published report on the relation between barberry and stripe rust epidemics.

There are nine different formae speciales (including *Pst*) of *Puccinia striiformis* depending on their host specificity and range. First, Erickson described five formae speciales; *Pst* on wheat, *P. striiformis* f. sp. *hordei* on barley, f. sp. *secales* on rye, f. sp. *elymi* on *Elymus* spp. and f. sp. agropyron on Aropyron spp. (Eriksson & Henning, 1896). In time, additional 4 formae speciales were stated by different publications: *P. striiformis* f. sp. *dactylidis* on orchard grass (*Dactylis glomerata*) (Manners, 1960; Tollenaar, 1967; Zadoks, 1961); *P. striiformis* f. sp. *poae* on Kentucky bluegrass (*Poa pratensis*) (Britton & Cummins, 1956; Murdoch et al., 1973; Tollenaar, 1967; Tollenaar & Houston, 1967); and *P. striiformis* f. sp. *leymi* on *Leymus secalinus* (Niu, Li, & Shang, 1991); *P. striiformis* f. sp. *pseudo-hordei* (Wellings et al., 2000; Wellings, 2011). Each formae speciale is specialized on their respective host; however, they may infect the other grass species if the environmental conditions and pathogen inoculum amount are favorable. Due to economic importance of wheat and barley, *P. striiformis* f. sp. *tritici* and *hordei* were studied extensively.

Pst isolates obtained from fields were analyzed for their virulence and avirulence attributes with the help of a defined set of wheat genotypes or single gene lines termed as wheat or yellow rust (YR) differential lines. Each race has different interaction with different host cultivars. The balance between virulence of the pathogen race and susceptibility of the wheat cultivar determines the impact of the disease. The emergence of new aggressive races causes broad range epidemics. Therefore, comprehending the disease progress along with resistance mechanism of the host in molecular biology level is key information for achieving success against wheat yellow rust disease. Fungicide may be a solution but deploying resistant cultivars is more environmentally friendly and cost-effective in long term.

1.3 Plant immunity

Plant species are under endless threat of phytopathogens since the beginning. Plants are lack of adaptive immunity unlike mammals, which is mobile and circulating through organism. However, they manage to survive against the pressure due to their innate immunity. Two defense layers against pathogen attacks are present in a plant cell. First, they have specific receptors to sense pathogen or microbe associated molecular patterns (PAMPs or MAMPs) which are called pattern recognition receptors (PRRs). PAMPs are patterns that are present in broad range of pathogen. For example, flagellin epitope (flg22) is a PAMP, which is recognized by a specific PRR named Flagellin sensing2 (FLS2) (Boller & Felix, 2009; Zipfel, 2009). There are also danger associated molecular patterns (DAMPs) such as polysaccharides of plant cell wall architecture (oligogalacturonides etc.). These DAMPs could also trigger similar response because their presence in a normal cell is not probable in natural conditions (Hückelhoven, 2007; Toyota et al., 2018). As PRRs detect PAMPs, PAMP triggered immunity (PTI) is stimulated. A series of phytohormone signaling is activated to counter pathogen attacks upon PTI activation. Cell wall thickening and lignification, ion fluxes and oxidative burst, release of reactive oxygen species, callose deposition, ethylene and salicylic production, stomatal closure etc. occur in response to PTI to halt pathogen growth and proliferation (Zipfel & Robatzek, 2010). The PTI response is a massive, repelling and intimidating shield against the pathogen menace. However, successful pathogens can evade, suppress or manipulate the PTI phenomenon with the aid of specific and small, proteinaceous compounds called effectors. Phytopathogen effectors could inhibit plant defense related enzymes, block or seize recognition of PAMPs by PRRs, jam the signaling system etc. (see Section 1.4) Plants need another defense strategy to protect themselves against the effectors. Hence, they possess a secondary and more specific system called effector triggered immunity (ETI) to meet the problem head on. In ETI, various cytoplasmic resistance (R) proteins are present in the plant cell in order to detect the effectors. These resistance (R) proteins share common features such as nucleotide binding (NB) domain and leucine rich region

(LRR), referred as NB-LRR or NLR occasionally due to resemblance of animal equivalents (Jones & Dangl, 2006). Generally, an R protein contains LRR region in their architecture for ligand recognition in the carboxyl-terminal (C-terminal) whereas nucleotide-binding domain remains at the center and is responsible for ATP/ADP binding (Bej et al., 2014; Kobe & Kajava, 2001). Moreover, in the amino-terminal (N-terminal), Toll/Interleukin-1 receptor (TIR) or coiled-coil (CC) domains could be existing (Peter N. Dodds & Rathjen, 2010). Based on their architecture, they could be categorized in two groups as TIR-NB-LRR and CC-NB-LRR. When the effector translocates in the host cell to contribute virulence, if the corresponding R protein is present, then interaction with the effector occurs. Now, the effector is termed as avirulence (Avr) protein because of perception accomplished. ETI response is stimulated and an array of secondary events occurs similar to PTI response. Except, ETI reaction may result in programmed cell death (PCD) to prevent further infection. To generalize, ETI and PTI are quite similar in the respect of triggered defense response. However, PTI is an extensive defense to broad range of pathogens including non-host and non-adaptive ones due to PAMP perception whereas ETI is an intensive defense against specific pathogens, which are adapted to host.

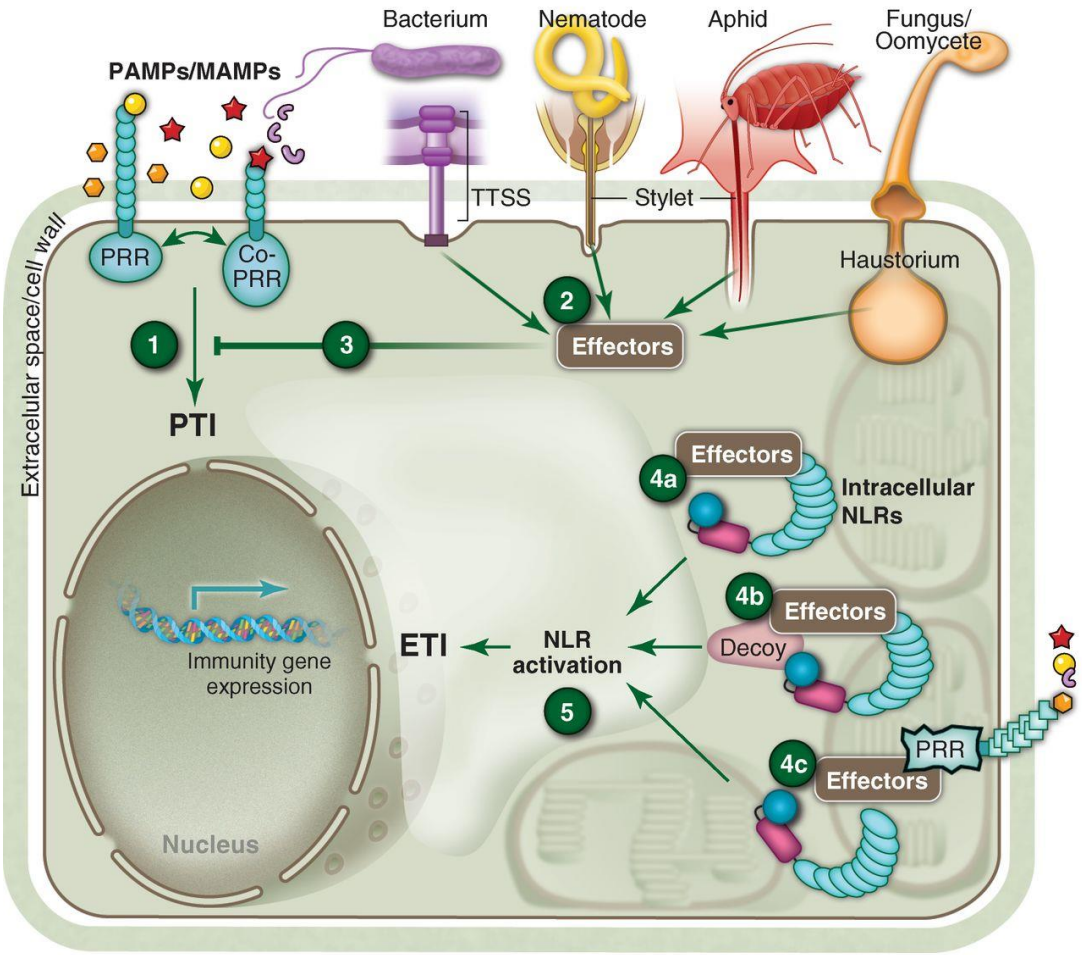


Figure 1.3. Schematic overview of plant innate immunity. (Dangl, Horvath, & Staskawich, 2013)

The recognition of the effectors by R proteins occurs either directly or indirectly. A certain number of R protein-effector interactions have been discovered in last decades. Each studied interaction raised new questions in plant innate immunity. For direct interaction, there is a specific Avr protein for each R gene product, described as gene-for-gene model by H. H. Flor in 1942 (Flor, 1971; Keen, 1990). Direct interaction of R proteins with the effectors has been demonstrated in several studies. For example; 1) AvrPita of rice blast fungus (*Magnaporthe grisea*) and Pi-ta (CC-NB-LRR type resistance protein) of rice (*Oryza sativa*) (Jia et al., 2000), 2) AvrPopP2 of bacterial wilt pathogen (*R. solanacearum*) and RRS1-R (TIR-NB-LRR type resistance protein)

of *Arabidopsis thaliana* (Deslandes et al., 2003), 3) AvrL567 of flax rust fungus (*Melampsora lini*) and L disease resistance protein (TIR-NB-LRR) of flax (*Linum usitatissimum*) (Dodds et al., 2006; Dodds et al., 2004) and 4) AvrM of flax rust fungus and M (TIR-NB-LRR type resistance protein) of flax (Catanzariti et al., 2006; Catanzariti et al., 2010). The indirect interaction has three different models and each covers the other's weaknesses. First model is the guard hypothesis. Avr proteins as an effector have virulence functions in the host cell which is lack of corresponding R protein. The effector protein modifies or acts on its target to achieve or enhance pathogenicity. The resistance arises from the recognition of the effector's target by R protein upon the modification. The indirect recognition of AvrPto and Pto kinase interaction by Prf (R protein) in tomato (Van Der Biezen & Jones, 1998) is explained by guard hypothesis in which Avr target is termed as 'guardee' and enhanced virulence occurs upon Avr targeting in the absence of R protein (Dangl & Jones, 2001). RIN4 protein of *A. thaliana* is a target for multiple pathogen effectors. AvrRpt2 attacks and cleaves RIN4, which is monitored by RPS2 resistance protein (Axtell & Staskawicz, 2003; Mackey et al., 2003). AvrB and AvrRPM1 phosphorylate RIN4 but this modification was detected by RPM1 (Mackey et al., 2002). Multiple Avr targeting of RIN4 and recognition by different R proteins is also described by guard hypothesis (Jones & Dangl, 2006). However, a 'guardee' is not evolutionary favorable in the plant population that is absent for R protein. Adaptation pressure drives guardee to evade interaction with the Avr in the R protein absent plants and vice versa in the plants that have R protein. The 'decoy model' is proposed to enlighten the evolutionary stress on 'guardee's (van der Hoorn & Kamoun, 2008). In the decoy model, Avr protein attributes its virulence function by manipulating target proteins in host cell, absent of R protein. Either by gene duplication or independent evolution of same function, a decoy protein is also present in host cell. The Avr acts on both real targets and the decoy proteins same time. However, resistance occurs when Avr modifies the decoy protein and this process is recognized by R protein. For instance, AvrPto targets kinase functions of FLS2 and CERK1 receptors as well as Pto kinase which is also similar to FLS2 and CERK1 (Gimenez-Ibanez et al., 2009; Zipfel & Rathjen, 2008). The decoy

model elucidates the tension of evolutionary force on target proteins in a beautiful manner, although; the function of the decoy is still unclear. The decoy is functional in resistance but it has no other function in the plant population that is lack of R protein. Pto is required for Prf activation and RIN4 has function in basal resistance (Kim et al., 2005; Liu et al., 2009; Marathe & Dinesh-Kumar, 2003; Mucyn et al., 2009). 'Bait and switch' model is a hybrid hypothesis that involved a cofactor protein as a bait for Avr protein, which is bound to R protein. Upon Avr-cofactor interaction, configuration of the R protein switch and down-stream process occurs (Collier & Moffett, 2009). In recent years, various studies about how R proteins work as hetero and homodimer complexes have been reported. For instance, AvrRps4 from *Pseudomonas syringae* and PopP2 from *Ralstonia solanacearum* are two Avr proteins that interact with RPS4 and RRS1 (TIR-NB-LRR type resistance proteins) of *A. thaliana* (Deslandes et al., 2003; Gassmann et al., 1999; Narusaka et al., 2009). Both AvrRps4 and PopP2 proteins bind WRKY domain of RRS1, which is bound to RPS4. As the recognition occurs, RPS4 is released from the heterodimer forms homodimer and leads to cell death (Williams et al., 2014). AvrPia and AvrPi-CO39 effectors bind to RATX1 domain of RGA5 that is also bound to RGA4 as a heterodimer complex (Cesari et al., 2013). RGA4 and RGA5 (CC-NB-LRR type resistance proteins) are required for resistance; expression of only RGA4 in *N. benthamiana* leaves triggered cell death (Césari et al., 2014). 'Integrated decoy' model was proposed to elucidate recent interaction studies (Cesari et al., 2014). In this model, target decoy protein is integrated to NB-LRR as a domain. Hence, direct and indirect interaction is coined up as a combined concept. All models are summarized as an illustration in Figure 1.4.

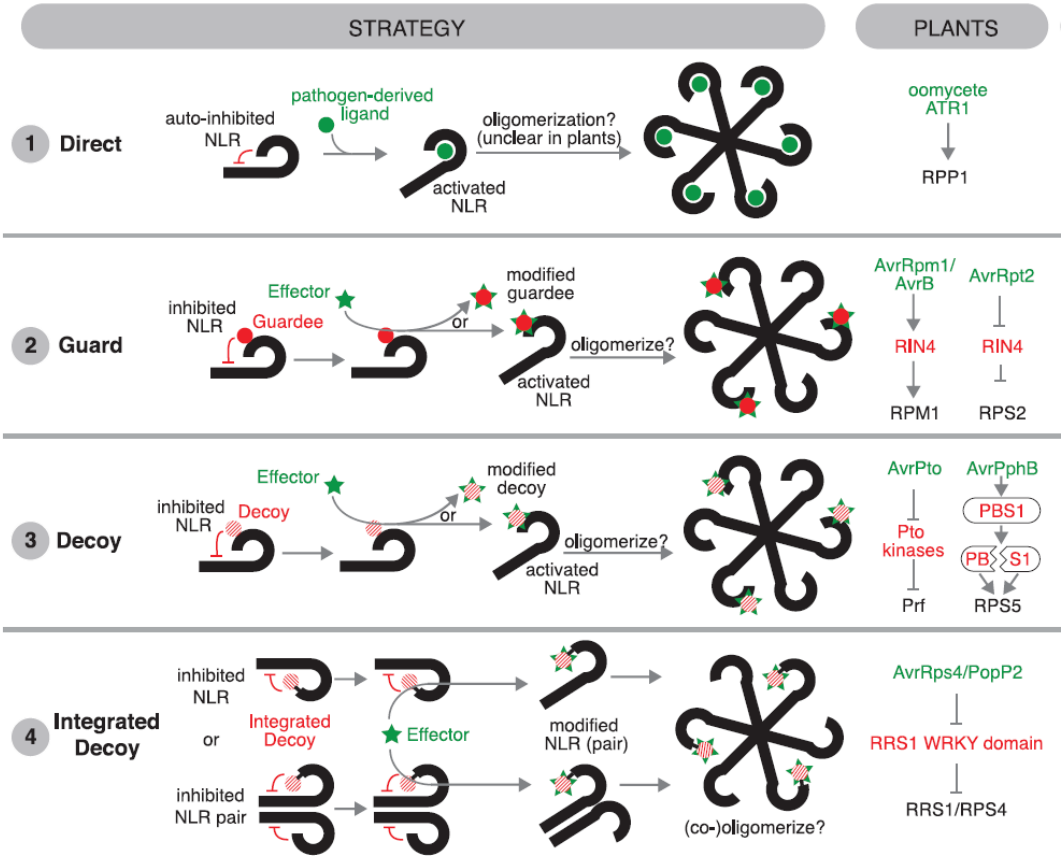


Figure 1.4. A graphical abstract to summarize current models used to elucidate Avr-R protein interactions in plant innate immunity.NLR: nucleotide binding, leucine rich (NB-LRR) proteins. 1) Direct interaction of Avr and R protein (gene-for-gene model). Indirect interaction of Avr and R protein 2) by monitoring changes on 'guardee' (guard model) and 3) modifications on 'decoy' protein (decoy model). 4) Combinatory approach of direct and indirect interaction; Avr acts on decoy domains of R protein (Integrated decoy model) (Jones et al., 2016).

The discovery of Avr and R proteins and their interaction mechanisms are crucial to solve the mystery of plant innate immunity in the scale of molecular biology. Each Avr and R protein helps us to obtain new proofs for existing models or proposing models that are more expedient to comprehend the resistance. Although both plant and animals have intrinsic R proteins (NB-LRR/NLR) that evolved from distinct derivatives of a common ancestral prokaryotic adenosine triphosphatase (ATPase), animals have additional complex immune system to survive (Jones et al., 2016). The

plants populate the planet relying on the innate immunity. We could help them flourish even more against diseases in hope for sustainable agriculture if we grasp the basics of plant innate immunity.

1.4 Effectors

Phytopathogen effectors are proteins, which are deployed to the apoplast or inside host cells to construct a favorable environment for pathogen growth and proliferation. Their roles may vary depending on each effector. For instance, their molecular and biological functions inside host may include suppression of immunity through interfering with PTI or ETI responses, modification of host target molecules to ensure pathogen fitness, jamming phytohormone signaling pathways or transcriptional regulators, enzymatic activity such as cell wall degrading enzymes (CWDEs) to promote growth, inhibition of host defense enzymes etc. (De Jonge et al., 2011). As we discussed in Section 1.3, the presence of an effector may contribute to pathogen virulence or avirulence depending on the interaction between host and pathogen. The host must recognize pathogen and its effectors to halt disease progress whereas pathogen must evade the recognition and manage to develop and proliferate with the aid of the effectors. Hence, the diversity and the extent of the pathogen's effector repertoire are crucial to pathogen survival and host adaptation. In this section, we briefly cover what are the common features and functions of phytopathogen effectors.

Studies on bacterial effectors are considerable in number. Our knowledge about the essentials of type 3 secretion system (T3SS) enables us to discover effectors from microbial genomes such as *Pseudomonas syringae* and *Xanthomonas* spp. (The Type 4 secretion system belongs to *Agrobacterium tumefaciens* will be covered later because they are important in delivery and integration of transfer DNA into host genome). On the other hand, the discovered effectors of filamentous pathogens (oomycetes and fungi) are reasonably inferior compared to T3SS effectors. Thus,

limited information is available for the effectors on filamentous fungi. Each effector untangled for its function and structure, is quite valuable for systematic characterization as well as its novelty. Filamentous fungi secrete the effectors via exocytosis at the site of hyphal tips toward apoplastic fluid upon host colonization. However; rusts, downy and powdery mildews deploy the effectors at the site of specialized feeding structure called 'haustoria' (Figure 1.5). The haustoria penetrate through host cell without disturbing the cell integrity and form a bulb like structure, which is the site for the effector production and feeding.

The location of the effector is illuminating its function. In general, there are two sites for effector location: apoplastic effectors or cytosolic effectors. Apoplastic effectors remain in the apoplastic fluid after secretion to achieve their functions such as CWDEs or necrosis and ethylene-inducing protein (NEP1) like proteins (NLPs) (Ottmann et al., 2009; Stotz et al., 2014). Small cysteine-rich effector proteins (SCEPs) are another type of apoplastic effectors; however, only limited information about their function is available so far. As the name suggests, they are rich in their cysteine amino acid number to form disulfide bridges which helps them for the stabilization of their tertiary structure in hostile environment of apoplast (Djonović et al., 2006; Stergiopoulos & de Wit, 2009). Through bioinformatics analysis, significant numbers of secreted proteins are predicted as SCEPs in both fungi and oomycetes (Saunders et al., 2012; Kim et al., 2016). Despite of tremendous efforts to elucidate their biological relevance in plant pathogen interaction, only few has been characterized hitherto. The SCEPs may inhibit the activity of host defense proteases such as Avr2, EPIC1 or EPIC2B (Tian et al., 2006; Shabab et al., 2008; van Esse et al., 2008). Co-operation between lysine motif (LysM)-containing chitin oligosaccharide elicitor-binding protein (CEBiP) and LysM-containing chitin elicitor receptor kinase-1 (OsCERK1) make the plant sense broad range of pathogens and activate chitin-triggered defense signaling (Kaku et al., 2006; Shimizu et al., 2010). Avr4 and Extracellular protein 6 (Ecp6) of *Cladosporium fulvum* have roles in evasion of the recognition of chitin oligosaccharides released from fungal cell wall due to host defense exochitinases.

Avr4 is a counter-defensive virulence factor that has a 'chitin-binding lectin containing an invertebrate chitin-binding domain' (CBM14). Hence, Avr4 could specifically bind to chitin oligomers not any other cell wall polysaccharides, in order to protect fungal hyphae against defensive plant chitinases (van den Burg et al., 2006). On the other hand, Ecp6 competes with host recognition receptors in binding chitin oligomers released as host chitinases attack (De Jonge et al., 2010). Cytoplasmic effectors are the second class of effectors based on their location of functionality. The cytoplasmic effectors of fungi need to be translocated inside host cell; however, their uptake mechanism is yet to be elucidated. The most studied entrance strategy exploits a conserved N-terminus motif of RxLR (Arg-x-Leu-Arg) followed by dEER (Asp-Glu-Glu-Arg) motif in the downstream of the oomycetes effectors. Avr3a and Avr1b effectors of *P. infestan* and *P. sojae* were validated for their entry inside the host cells using stable tranble transformants (Dou et al., 2008; Whisson et al., 2007). The RxLR motif was reported to interact with the phospholipid phosphatidylinositol-3-phosphate (PI3P) at the cell surface of the plant cell, thereby the effector translocates into the host cytoplasm in pathogen independent manner (Kale et al., 2010). Additionally, there are reports claiming the presence of RxLR-like motifs is sufficient to accomplish translocation inside host cells (Kale et al., 2010; Plett et al., 2011; Rafiqi et al., 2010). ToxA effector was previously identified to induce cell death in the fungal pathogens of wheat (*Pyrenophora tritici repentis* and *Stagonospora nodorum*) by separate studies (Liu et al., 2006; Manning and Ciuffetti, 2005). RGD (Arg-Gly-Asp) motif is suspected to involve in internalization of the ToxA inside the host cell (Manning et al., 2008). The effectors belong to Crinkler (CRN) family were discovered in oomycete pathogen, *P. infestans* (Torto et al., 2003). The CRN effectors were able to induce cell death and crinkling upon activation of defense genes in host and possess LFLAK motif in the N-terminus (Haas et al., 2009; Torto et al., 2003). The chimeric protein consisting of the fusion between N-terminus CRN of AeCRN5 (*Aphanomyces euteiches*) (or CRN3, CRN8 and CRN16) and C-terminus of Avr3a (*P. infestans*) was able to translocate into *N. benthamiana* cells, showing LFLAK is sufficient as RxLR in terms of cell entry (Schornack et al., 2010). CHxC (Cys-His-x-Cys) motif was

identified in several predicted secreted proteins of *Albugo laibachii* and CHxC9 was demonstrated to target C-terminus of Avr3a into host cells as an N-terminus fusion (Kemen et al., 2011). Unlike oomycete effectors, the fungal effectors are lack of conserved motifs. There are very few reports for the presence of conserved motifs and they are not studied for their function in translocation. The effectors of rice blast were reported to bear a conserved [LI]xAR (Leu/Ile-x-Ala-Arg) motif (Yoshida et al., 2009). YxSL[RK] motif of *Pythium ultimum* and [RK]VY[LI]R motif of *Blumeria graminis* f.sp. *hordei* were discovered as candidate conserved motifs (Lévesque et al., 2010; Ridout et al., 2006). The most notably, [FYW]xC motif was identified in number of candidate effectors of powdery mildew and rusts (Godfrey et al., 2010). However, its significance is yet to be clarified. It is possible that the haustoria provide a handy interface with enough proximity for the effector translocation in pathogen dependent or independent manner (Figure 1.5). The biotrophic interfacial complex (BIC) of *Magnaporthe oryzae,* was demonstrated for involvement in cytoplasmic effector production unlike apoplastic ones (Khang et al., 2010). In the same fashion, it is reasonable to speculate haustoria could manage such supply. With internal motifs or pathogen dependent manner, effectors are pursuing their targets inside host cytoplasm in order to achieve conditions favorable to pathogen growth. Too little is known for cytoplasmic effectors, although, they can reprogram transcription, suppress defense, jam signals, provide nutrition and so on.

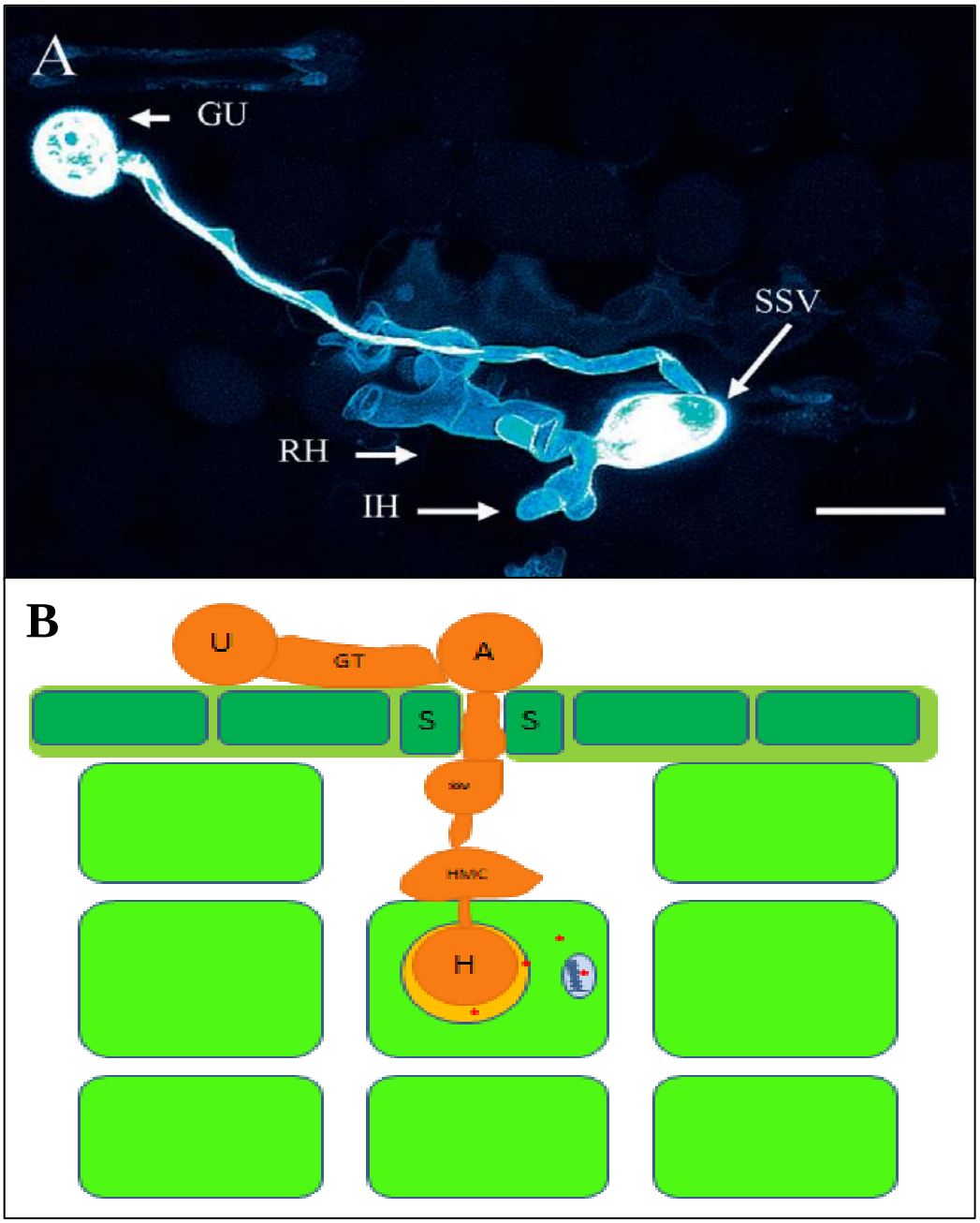


Figure 1.5. Production, delivery and trafficking of effectors via haustoria (H) interface of Pst upon infection. S: Stomata cells, A: Apressorrium, GT: Germ tube, U: Urediospore, RH: Runner hypae, IH: Infection Hyphae, HMC: Haustorial mother cell, Red stars: Effectors. A) Real image (Bozkurt et al., 2010) and B) Schematic representation of the haustoria formation during infection.

1.4.1 *Pst* effectors

There is only small number of *Pst* effectors known to study up to now. Lack of common or conserved features in *Pst* effectors as aforementioned in other pathogens, halt the investigation. Moreover, *Pst* is not culturable in any known growth medium and the wheat is not easy to work with either. Hence, the studies on *Pst* effectors are limited in case of screening candidate effectors. The list of published *Pst* effectors hitherto, are briefly introduced in this section.

Ps87 is an effector candidates discovered in the cDNA library of germinated urediospores of PstCY32 (Zhang et al., 2008). Ps87 was reported to bear RxLR like motif that successfully directs the effector inside the host cell (Gu et al., 2011). PEC6 (*Pst* effector candidate 6) effector was identified to interact with adenine kinase in host cells in order to suppress PTI by hindering reactive oxygen species (ROS) release and obstructing callose deposition (Liu et al., 2016). NPR1 (Non-expresser of PR genes 1) is known to involve in transcriptional regulation of elements related to pathogenesis related (PR) gene action in the course of pathogen exposure (Cao et al., 1994; Mou et al., 2003). PNPi (Puccinia NPR1 interactor) have DPBB-1 domain to interact with NPR1 in the nucleus to jam its interaction with the corresponding transcription factor of defense genes (Wang et al., 2016). PstHa5a23 is one of the candidate effectors that identified in the haustorial cDNA library of *Pst*-78 (Yin et al., 2009). PstHa5a23 was studied using sequences of its full-length homolog Pstg00676 (Pst-78) in both wheat and *N. benthamiana* systems (Cheng et al., 2017). The authors discovered PstHa5a23 targets host cytoplasm and it suppresses the cell death triggered by several elicitors (INF1, BAX1, MKK1, and NPK1) in heterologous systems (See Section 1.5.5). PstSCR1 (previously PstHa2a5) was reported to trigger severe cell death upon translocation into the apoplastic fluid whereas the effector without its signal peptide (SP) sequence was localized in the cytoplasm showed no symptoms at all (Dagvadorj et al., 2017). A number of *Pst* effector candidates were investigated to pinpoint their interactors inside host cells of *N. benthamiana* leaves (Petre et al., 2016).

Similarly, nine *Pst* effector candidates along with 11 Pgt effector candidates were studied and seven of the *Pst* effectors were reported to suppress cell death (Ramachandran et al., 2017). Recently, an effector candidate (Pst8713) of Pst-CY32 was publicized for its ability to suppress the cell death triggered by INF1 and BAX hampering ROS and callose deposition (Zhao et al., 2018). Moreover, the authors demonstrated the effector have certain involvement in virulence by means of reduction in uredinia production (Zhao et al., 2018).

Table 1.2. The list of Pst effectors reported in literature.

Effector	Length	Location	Interactor	Function	References
PEC6	88	Cytoplasm and Nucleus	ADK1	Suppressor of PTI	Liu et al., 2016
PNPi	333	Cytoplasm and Nucleus	NPR1	Suppressor of defense responses	Wang et al., 2016
PstHa5a23	108	Cytoplasm	-	Suppressor of PTI and cell death, enhanced virulence	Cheng et al., 2017
PstSCR1	116	Apoplast	-	Cell death elicitor	Dagvadorj et al., 2017
Pst8713	114	Cytoplasm and Nucleus	-	Suppressor of PTI and cell death	Zhao et al., 2018
Ps87	85	Cytoplasm	-	No effect	Gu et al., 2011
Shr1 (Pstg00494)	199			Suppressor of cell death	Ramachandran et al., 2017
Shr2 (Pstg01062)	182			Suppressor of cell death	Ramachandran et al., 2017
Shr3 (Pstg01724)	114			Suppressor of cell death	Ramachandran et al., 2017
Shr4 (Pstg09266)	191			Suppressor of cell death	Ramachandran et al., 2017
Shr5 (Pstg10812)	105			Suppressor of cell death	Ramachandran et al., 2017
Shr6 (Pstg14250)	199			Suppressor of cell death	Ramachandran et al., 2017

Shr7 (Pstg14695)	151			Suppressor of cell death, PTI and HR	Ramachandran et al., 2017
Pst02549	297	P bodies	EDC4	-	Petre et al., 2016
Pst18220	110	Chloroplast and Nucleus	ABC transporter F family member 4, THO complex subunit 2, DNA damage binding protein 1	-	Petre et al., 2016
Pst03196	206	Chloroplast	-	-	Petre et al., 2016
Pst05023	281	Endomembrane	RNA recognition motif containing protein	-	Petre et al., 2016
Pst05258	256	Cytoplasm and Nucleus	-	-	Petre et al., 2016
Pst05006	201	Cytoplasm and Nucleus		-	Petre et al., 2016
Pst05302	160	Cytoplasm and Nucleus		-	Petre et al., 2016
Pst08468	206	Cytoplasm and Nucleus	SNF4, SNF4 like protein, SNF1 related protein kinase	-	Petre et al., 2016
Pst11721	250	Nucleus	Chaperonin, Serine/threonine-protein phosphatase 2A, Dihydrodipicolinate reductase 3	-	Petre et al., 2016
Pst18447	146	Nucleus	-	-	Petre et al., 2016
Pst15391	256	Nucleus	-	-	Petre et al., 2016
Pst10977	171	Cytoplasm and Nucleus	-		Petre et al., 2016
Pst12160	168	Cytoplasm and Nucleus	Signal recognition particle 54 kDa protein, oxidoreductase, Ubiquitin 1	-	Petre et al., 2016
Pst15642	102	Cytoplasm and Nucleus	-	-	Petre et al., 2016

Pst18221	112	Cytoplasm and Nucleus	-	-	Petre et al., 2016
Pst15964	128	Cytoplasm and Nucleus	-	-	Petre et al., 2016

1.5 Strategies for the discovery and the investigation of the effectors

The biological importance of effectors as virulence factors and avirulence determinants commence the new era of effectoromics. Common features of the effectors such as secretion to apoplast or host cell enable for high through put discovery of effectors. Unique features of the effectors compel the studies to elucidate their functions. Therefore, there is considerable number of methods in effectoromics research including *in silico, in vitro* and *in vivo* approaches. Sequencing studies increase the number of data related pathogen races, disease progress and so on. *In silico* processes aid in evaluating, dissecting and filtering the gigantic datasets generated through sequencing. The most laborious and time-consuming part is to validate each effector candidate for their function using *in planta* and *in vitro* practices. In this section, the state-of-the-art strategies in phytopathogen effector research will be discussed briefly.

1.5.1 Genomics, transcriptomics and proteomics

Improvements in next generation sequencing (NGS) enable the researchers for generating genomics data on much more complicated and challenging organisms. Moreover, decline in the cost of the sequencing services leads to acquire more data on populations, isolates and races of a particular organism. The significant outcome of this trend is the ability to make comparisons between different organisms, isolates in an attempt to discover virulence and avirulence factors. For instance, genome wide association analysis and correlation analysis could be applicable with increasing

datasets. Similarly, transcripts profile of a pathogen could be monitored during different time intervals and various environmental conditions during the infection process thanks to adaptability and accessibility of NGS.

Obligate biotrophic nature of *Pst* constrains the studies on the yellow rust disease. *Pst* is not culturable on artificial media but can only be maintained as urediniospores on living host. First identification of functional genes in disease was studied on urediniospores of *Pst*-78 using full-length cDNA library (Ling et al., 2007). Gene expression analysis was reported on germinated urediniospores of *Pst*-CY32 using ESTs (Zhang et al., 2008). Haustoria specific genes were identified on *Pst*-78 cDNA library and authors defined secreted proteins that are rich in haustoria and various infection stages (Yin et al., 2009). The genome of *Pst*-130 was sequenced using NGS technology (Cantu et al., 2011). Zheng et al. sequenced Chinese isolate *Pst*CYR-32 (09-001) and analyze the origin of the isolate by comparing four *Pst* isolates (*Pst*-CY23, 104E137A, PK-CDRD, Hu09-2) from different geographical region (Zheng et al., 2013). *Pst*-78 (2K-041) genome was published along with detailed comparative analysis between *Pgt* and *Ptt* (Cuomo et al., 2017). Broad Institute released genome sequences of *Pst*-78 (2K-041), *Pst*-1 (3-5-79), *Pst*-127 (08-220), *Pst*CYR-32 (09-001) (Cuomo et al., 2017; Xia et al., 2017; Zheng et al., 2013). In addition, genome sequences of *Puccinia graminis*, *Pgt* (CRL 75-36-700-3) and *Puccinia triticina*, *Ptt* (BBDD) was published as a publicly available reference dataset, which is also useful for comparison studies (http://www.broadinstitute.org/) (Cuomo et al., 2017; Duplessis et al., 2011). The genome sequences of the four races, including *Pst*-87/7, *Pst*-08/21 (two United Kingdom (UK) races), *Pst*-21, *Pst*-43 (two United States (US) races), were reported in a publication by Cantu and his colleagues. Additionally, the gene expression data belong to different time points (6 and 14 days post inoculations (dpi)) of the infection were described as well as haustoria specific genes (Cantu et al., 2013). Hubbard et al. conducted a survey on *Pst* isolates collected in UK fields in 2013 using RNAseq approach (Hubbard et al., 2015). The authors investigated the evolutionary resemblance of harvested *Pst* isolates to the historical *Pst* isolates (14

UK and 7 French isolates) and 6 additional *Pst* isolate through the whole genome sequencing (WGS). They discovered that the field isolates were not related to old isolates but they were possibly originated from foreign *Pst* population. Garnica et al. reported RNAseq data generated from both haustoria and germinated urediospores of an Australian isolate (104E137A) (Garnica et al., 2013). A custom-made microarray chip was developed to reveal expression profile of suspected genes obtained from past reports (Huang et al., 2011). In a recent study, seven new races of *Pst* was sequenced using NGS and combined with seven older published genomes. Total 14 races of *Pst* was subjected to correlation analysis in order to point out Avr candidates (Xia et al., 2017). A study about proteome profile of compatible interaction between wheat and *Pst* revealed the evidences about proteins involved in disease process (Demirci et al., 2016). Another proteome study listed proteins of *Pst* that are active in urediospores and germtubes, using iTRAQ (isobaric tag for relative and absolute quantitation) method and qRT-PCR for validation (Zhao et al., 2016). Alterations in the proteome content of *Pst* urediospores in response to application of UV-B radiation were reported for three different Chinese race (CYR31, CYR32, CYR33) to elucidate deviations in virulence mechanism (Zhao et al., 2018).

In summary, we discussed the published sequence data that is available about *Pst*. To provide better perspective, Table 1.3 was organized using data reported in literature that comes to our knowledge. It is clear that frequent number of studies were conducted with the help of different strategies of genomics, transcriptomics and proteomics; even though, *Pst* is troublesome to study due to its obligate biotrophic lifestyle. Destructive impact of the yellow rust disease leads each country to survey the pathogen to protect wheat fields. However, Turkey progress at modest level related to this aspect. The country that is the evolutionary home for wheat, should act fast and strong to secure food chain by increasing the research on *Pst*. Genome sequencing is laborious and expensive. Proteome studies generate a narrow range of information because of the abundancy problem of pathogen proteins. Transcriptome sequencing is a fast and cost-effective action to initiate researches on Turkish *Pst* races. Microarray

profiling is still useful but slight expensive. Another disadvantage is that only a set of genes with known sequences could be monitored through microarray studies not the novel ones.

Table 1.3. List of the 'Omics' studies conducted on Pst.

Genome Sequencing Studies			
Isolate/Race		Origin	Source
Pst-78 (2K-041)	PSTv-35	US	Cuomo et al., 2017
Pst-130 (07-168)	PSTv-69	US	Cantu et al., 2011
Pst-CYR32 (09-001)	PSTv-37	China	Zheng et al., 2013
Pst-CY23		China	Zheng et al., 2013
104E137A		Australia	Zheng et al., 2013
PK-CDRD		Pakistan	Zheng et al., 2013
Hu09-2		Hungary	Zheng et al., 2013
Pst-1 (3-5-79)	PSTv-1	US	Broad Institute
Pst-21 (07-214)	PSTv-1	US	Cantu et al., 2013
Pst-43 (03-338)	PSTv-27	US	Cantu et al., 2013
Pst-127 (08-220)	PSTv-11	US	Broad Institute
PST-87/7		UK	Cantu et al., 2013
PST-08/21		UK	Cantu et al., 2013
14 UK isolates and 7 French isolates.		UK and France (21 Historical isolates)	Hubbard et al., 2015
PST-78/66, PST-12/86, PST-12/83, PST-11/13, PST-11/128 PST-11/08			Hubbard et al., 2015
11-281	PSTv-18	US	Xia et al., 2017
Pst-127 (09-134)	PSTv-11	US	Xia et al., 2017
12-248	PSTv-2	US	Xia et al., 2017
12-346	PSTv-40	US	Xia et al., 2017
12-368	PSTv-4	US	Xia et al., 2017
PK08-2	PSTv-119	Pakistan	Xia et al., 2017
841541:430	360E137A	Australia	Xia et al., 2017
Transcriptome Sequencing and Microarray Studies			
Isolate/Race	Strategy	Sample Conditions	Source
Pst-78 (2K-041)	cDNA library (EST)	Urediniospores	Ling et al., 2007
Pst-CYR32 (09-001)	cDNA library (full length)	Germinated urediniospores	Zhang et al., 2008

Pst-78 (2K-041)	cDNA library (EST)	Haustoria	Yin et al., 2009
PST-08/21	RNAseq	6 and 14 dpi infected wheat samples and haustoria	Cantu et al., 2013
Pst-78 (2K-041)	Microarray	12, 24 and 48 hpi, 7 and 14 dpi infected wheat samples and germinated urediniospore	Huang et al., 2011
104E137A	RNAseq	Haustoria and germinated urediospores	Garnica et al., 2013
PstMix (UK field isolates in 2013)	RNAseq	Infected leaves	Hubbard et al., 2015
PST-78/66, PST-12/86, PST-12/83, PST-11/13, PST-11/128 PST-11/08	RNAseq	Infected leaves	Hubbard et al., 2015
Proteome Sequencing Studies			
Isolate/Race	Strategy	Sample Condition	Source
PstMix (Turkey field isolates in 2010)	Nano LC-ESI-MS/MS	1, 2, 3, 4 dpi infected leaves	Demirci et al., 2016
Pst-CYR32	MALDI-TOF/TOF tandem MS and iTRAQ	Urediniospores and germinated urediospores	Zhao et al., 2016
Pst-CYR31, Pst-CYR32, Pst-CYR33	Nano LC-ESI-MS/MS and iTRAQ	Urediospores and UV-B radiation applied urediospores	Zhao et al., 2018

The pace of effector evolution and the emergence of new races lead the amassed number of suspects in effector biology. Each generated data of 'omics' related to *Pst*, aids us to discover, compare and pinpoint direct and indirect players in disease formation. Hence, it is crucial to increase the knowledge about different races, time points of the disease and environmental conditions. In this section, we covered the previous works about *Pst* using the strategy of genomics, transcriptomics and proteomics. Moreover, we emphasize the effectiveness of omics technologies in providing bulk data on various races and different phases of the disease in hope to screen molecular evidences. The data itself is beneficial to understand the disease;

nevertheless, they are also key resource for data mining approaches due to improvement in the bioinformatics analysis strategies.

1.5.2 Data mining

Data mining is a popular terminology to explain the studies conducted on big data sets using statistics, prediction algorithms, deep machine learning systems (artificial intelligence) and informatics in order to evaluate outcomes, to pinpoint crucial subset of data from bulk collection, to predict future patterns. However, in this section, we will use data mining as a terminology to cover all *in silico* strategies for handling and characterizing bulk data generated from various sequencing strategies with the intention to dissect most relevant information by the aid of bioinformatics. Prediction programs, computational methods and machine learning algorithms to predict and characterize effectors of phytopathogens will be discussed to introduce the application of *in silico* analysis on the studies of the effector biology.

In the previous section, the efficiency of 'Omics' studies were explained to disclose vital proteins of phytopathogens. Yet, the datasets obtained through sequencing are quite substantial. Furthermore, testing the uncovered number of genes is laborious, time-consuming and costly. There is a need for pooling down the most probable sets of candidates so that they can be experimentally tested for function. Data mining is useful strategy to narrow down candidate effectors. It practices our prior knowledge about effectors to predict new candidates. For instance, we know that secreted proteins are important in achieving virulence. Hence, predicting the secreted proteome or 'secretome' catalogue of a pathogen is a worthy way to start. Of course, the predicted subset may not be secreted or be irrelevant to virulence but we are now scanning on smaller list. Each prediction and characterization will increase the success of trials. The first attempt to dissect secretome was reported on haustorial cDNA library of *Pst*-78 (Yin et al., 2009). We (as Akkaya Research Group) studied a number of candidate

effectors of this publication which will be discussed later (Dagvadorj et al., 2017) (Andac et al., unpublished data), (Özketen A. Ç., M.Sc Thesis 2013). Duplessis et al., published genome sequences of poplar leaf rust *Melampsora larici populina* (*Mlp*) and wheat stem rust *P. graminis* f. sp *tritici* (*Pgt*) (Duplessis et al., 2011). Moreover, the group predicted the secretome of the pathogens and small-secreted proteins (SSPs). A pipeline to discover and characterize candidate effector proteins were defined in hierarchical clustering study using same data of pathogens; *Mlp* and *Pgt* (Saunders et al., 2012). Subsequently, the secretome and effectorome era has begun for phytopathogens. Numerous number of data generated with genomics, transcriptomics and proteomics were subjected to secretome prediction and characterization by several researchers (Cantu et al., 2011, 2013; Cuomo et al., 2017; Demirci et al., 2016; Garnica et al., 2013; Huang et al., 2011; Xia et al., 2016a; Xia et al., 2016b; Xia et al., 2017; Zheng et al., 2013).

The basis of secretome prediction depends on two rule: i) presence of secretion signal and ii) absence of transmembrane helices. A protein could be secreted either by classical or non-classical pathway. In the classical pathway, presence of N-terminus secretion signal or signal peptide is required for translocation through endoplasmic reticulum/golgi dependent secretory pathway (Nickel, 2003). Non-classical pathway is lack of any secretion signal contradictory to conventional path (Stein et al., 2014). However, secretome prediction is conducted frequently based on classical secretion even though some proteins follow non-classical pathways. Absence of any transmembrane helices is important to rule out any membrane-destined protein. After a secretome is defined, candidate effectors are predicted by means of certain parameters established on known effectors. Effector proteins are generally short in length. Some apoplastic effectors are rich in their cysteine content to provide more stability in the hostile environment of apoplastic fluid. Conserved motifs were also detected in amino acid sequence of fungal effectors as we discussed in Section 1.4. Effectors generally show no homology to known domains except the ones associated with pathogenicity. Effectors could be encoded by genes with long intergenic regions

and they may contain internal repeats. Hence, a researcher could set indefinite number of pipelines for effector mining using different filtering parameters based on known effector functions. A well-accepted pipeline was defined by Saunders et al. to pinpoint candidate secreted effector proteins (CSEPs) of fungal pathogens (Duplessis et al., 2011) (Saunders et al., 2012).

Bioinformatics is advantageous for *in silico* predictions of effectors. The discovery of each fresh effector and advancement in technology offers new programs for prediction. A small list of programs and databases convenient for effector discovery and characterization was presented as a Table 1.4. Each program uses different strategy such as sequence similarity, biochemical nature of its composition, presence of known signals and sequences for diverse sets of tasks including subcellular localization prediction, conserved domain discovery, structure and function deduction. Among these strategies, machine learning is new and it is recently introduced to effector prediction. Algorithms compare and learn experimentally validated sets of positive and negative results in order to forecast a novel protein belongs to which group. For example, EffectorP is the first reported machine-learning program to predict effectors from other secreted proteins (Sperschneider et al., 2016). EffectorP 2.0 has upgrade for increased accuracy released in advanced (Sperschneider et al., 2018a). ApoplastP and Localizer are other programs to calculate the subcellular localization of an effector inside/outside host plant (Sperschneider et al., 2017; Sperschneider et al., 2018b).

Table 1.4. Valuable databases and prediction programs used in effector mining.

Databases and Annotation Programs			
Name	Definition	Source (Web Link)	Reference
PHI-base	'Pathogen Host Interaction database' provides information about experimentally verified virulence and pathogenesis associated genes of pathogens infecting various host organisms (plant, fungi, human etc.)	www.phi-base.org/	(Urban et al., 2015) (Urban et al., 2017)
EffectorDB	Database for predicted Rare orthologous groups/Lateral gene transfer groups in fungal effectors	effectordb.com/	
FSD	'Fungal secretome database' generated using the data of fungal species stored on 'Comparative Fungal Genomics Platform'	fsd.snu.ac.kr/	(Choi et al., 2010)
DFVF	'Database of virulence factors in fungal pathogens' constructed via a novel prediction algorithm	sysbio.unl.edu/DFVF/	(Lu et al., 2012)
FunSecKB	'Fungal Secretome Knowledge Base' contains predicted fungal secreted proteins extracted from NCBI RefSeq	bioinformatics.ysu.edu/secretomes/fungi.php	(Lum & Min, 2011)
MEROPS	Database for peptidases, inhibitors and substrates of peptidases	www.ebi.ac.uk/merops/	(Rawlings et al., 2014, Rawlings et al., 2018)
dbCAN	Database of 'Carbohydrate active enzyme' (CAZymes) for automated CAZymes Annotation	cys.bios.niu.edu/dbCAN/	(Yin et al., 2012)
CAZy	CAZymes database of enzymes that degrade, modify, or create glycosidic bonds	www.cazy.org/	(Lombard et al., 2014)
LED	'Lipase engineering database' holds information about Lipases and Lipase associated proteins	www.led.uni-stuttgart.de/	(Fischer & Pleiss, 2003)
fPoxDB	'Fungal peroxidase database' fabricated via a novel prediction platform on fungal genomes	peroxidase.riceblast.snu.ac.kr/	(Choi et al., 2014)
Blast2GO	Gene annotation, visualization and analysis tool	www.blast2go.com/	(Conesa et al., 2005)
BlastKOALA	Gene annotation and characterization tool	www.kegg.jp/blastkoala/	(Kanehisa et al., 2016)

AmiGO	Gene ontology and annotation database	amigo.geneontology.org/amigo	(Carbon et al., 2009)
PFAM	Protein families database	https://pfam.xfam.org/	(Finn et al., 2014, 2016)
Smart	'Simple Modular Architecture Research Tool' for detection of protein domains and domain architectures	smart.embl-heidelberg.de/	(Letunic & Bork, 2018; Letunic et al., 2015)
CDD	'Conserved domain database' and domain analysis tool	https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi	(Marchler-Bauer et al., 2017, 2015, 2011)
Galaxy	An open-source, web-based platform designed with a user-friendly interface for researchers, could be used for effector mining	usegalaxy.org/ galaxyproject.org/	(Cock & Pritchard, 2014)
Prediction Programs			
SignalP	N-terminus secretion Signal Prediction based on artificial neural networks	www.cbs.dtu.dk/services/SignalP/	(Petersen et al., 2011)
SecretomeP	*Ab initio* prediction of non-classical protein secretion	www.cbs.dtu.dk/services/SecretomeP/	(Bendtsen et al., 2004)
TargetP	Prediction of subcellular localization of proteins	www.cbs.dtu.dk/services/TargetP/	(Emanuelsson et al., 2000)
EffectorP	Fungal effector prediction via machine learning method	http://effectorp.csiro.au/	(Sperschneider et al., 2016, 2018a)
ApoplastP	Prediction of proteins localized into apoplast region of plant via machine learning method	http://apoplastp.csiro.au/	(Sperschneider et al., 2017)
TMHMM	Prediction of transmembrane helices of proteins	www.cbs.dtu.dk/services/TMHMM/	(Krogh et al., 2001)
WolfPSORT	Prediction of subcellular localization of proteins	www.genscript.com/wolf-psort.html https://wolfpsort.hgc.jp/	(Horton et al., 2007)
Cello	Prediction of subcellular localization of proteins	cello.life.nctu.edu.tw/	(Yu et al., 2006)
Cello2GO	Prediction of subcellular localization of proteins with annotation	cello.life.nctu.edu.tw/cello2go/	(Yu et al., 2014)
Localizer	Prediction of subcellular localization of plant and effector proteins via machine learning	localizer.csiro.au/	(Sperschneider et al., 2018b)
Secretool	Prediction and characterization of fungal secreted proteins	genomics.cicbiogune.es/SECRETOOL/	(Cortázar et al., 2014)
PSIPRED	Secondary structure prediction and various analysis	bioinf.cs.ucl.ac.uk/psipred/	(Buchan et al., 2013; Jones et al., 1999)

MEME Suite	Motif discovery, enrichment and analysis tools	meme-suite.org/	(Bailey et al., 2009; Bailey & Elkan, 1994)

Data mining enable the researchers to filter amassed numbers of proteins in terms of probability of relevance. For instance, if a scientist wants to study effectors that have peptidase like function in virulence, they could sort out secretome repertoire by monitoring peptidase domains. Undeniably, prediction does not mean that the sorted sets of proteins will always have peptidase function. Subsequently, they should be verified experimentally. Likewise, data mining offers candidate effectors for functional verification. Each different pipeline yields different list of candidate effectors. Hence, each catalogs of candidates holds false positives and neglected false negatives. However, the advantages of mitigation of large datasets are greater than disadvantages. Fungal effectors do not share conserved sequence motifs, sequence similarity and common features in broad-spectrum. Hence, the improvements in data mining will ease up the studies on effector biology.

1.5.3 Type 3 secretion system based effector screening

The challenging part for effector investigation is discovery and experimental validation of its function. The host of *Pst* is wheat, which has hard to penetrate cuticle layer. The transformation of wheat is not easy. Horizontal gene via *Agrobacterium tumefaciens* mediated gene transfer is not feasible in wheat. On the other hand, particular reports claimed *A. tumefaciens* LBA4404 and COR308 able to achieve gene transfer on wheat cultivar, Thatcher (Cuomo et al., 2017; Panwar et al., 2013). We as Akkaya research group worked with same *A. tumefaciens* strains (LBA4404 and COR308) to express a particular effector under the cauliflower mosaic virus (CaMV) 35S promoter on AvocetS, yet we observed no expression (Ayşe Andaç, data unpublished). The other method is to use biolistics or gene bombardments for gene expression. Recently, a bacterial delivery system is engineered as the Effector-to-Host Analyzer (EtHAn) system through harnessing Type 3 secretion of *Pseudomonas syringae pv. syringae 61* for effector delivery by stably integrating the hrp/hrc region into the genome of *Pseudomonas fluorescens Pf0-1* (Thomas et al., 2009). pEDV6 Gateway destination vector (Appendix A) was constructed manipulating N-terminal amino acid sequence of AvrRPS4 for type 3 secretion of any effector of interest cloned into vector (Sohn et al., 2007). The capability of pEDV6 vector to deliver effectors into wheat and barley cultivars was studied extensively using *Pseudomonas syringae* DC3000 and *P. fluorescens* EtHAn; although, the latter was observed to awake less basal symptoms on experimented wheat cultivars (Yin & Hulbert, 2011). An effector of *P. graminis* f.sp *tritici* (*Pgt*AUSPE-10-1) triggered a hypersensitive response (HR) of cell death on wheat cultivar W3534 (Sr22+) using *P. fluorescens* EtHAn in genotype specific manner (Upadhyaya et al., 2014). However, *Pgt*AUSPE-10-1 did not produce a HR on other Sr22 carrying wheat cultivar neither the effector produced as a fusion of AvrRPS4 unlike AvrRPM1 fusion (Upadhyaya et al., 2014). An effector candidate of *P. striiformis* f.sp. *tritici* (*Pst*PEC6) expressed by *P. fluorescens* EtHAn mediated type 3 secretion system (T3SS) was able to suppress PTI and allow induction of HR by *P. syringae* DC3000 on wheat (Liu et al., 2016). On the contrary, *Pst*Shr7

significantly suppressed the HR symptoms elicited by *P. syringae* DC3000 infiltration on wheat cultivar Nugaines (Ramachandran et al., 2017). The conflict between two reports arises from the different results generated on co-infiltrated regions of both *P. fluorescens* EtHAn and *P. syringae* DC3000. Liu et al. observed no symptoms on co-infiltrated area similar to PTI elicitation-supression assays of subsequent *P. fluorescens* EtHAn and *P. syringae* DC3000 challenge on *N. benthamiana* (Chakravarthy et al., 2009). In contrast, Ramachandran et al. witnessed a cell death in co-infiltrated wheat leaves. These results suggest that both effector detector vectors (pEDV) and Effector host analyzer (EtHAn) systems are beneficial to investigate functions of the effectors in homologous and heterologous systems. However, a sure-footed attitude is required to prevent controversial and false positive outcomes.

1.5.4 Host induced gene silencing

Gene silencing strategy is an effective tactic to investigate biological significance of candidate genes. The gene silencing is possible by means of various approaches to deliver silencing constructs of antisense RNA. For instance, virus based elements could be engineered for expression of antisense RNA in host plants as in virus induced gene silencing (VIGS). VIGS is relatively straightforward and trouble-free in functional genomics of plants. However, this is not the case for plant pathogens. Particularly the obligate biotrophic pathogens are inaccessible for silencing constructs in many occasions. Host induced gene silencing (HIGS) exploits the ability of double-stranded RNA (dsRNA) or small interfering RNA (siRNA) translocation from host into pathogen (Nowara et al., 2010). Nowara et al. validated that HIGS appears in *Blumeria graminis* (an obligate biotrophic pathogen of barley) upon expression of hairpin RNA or dsRNA constructs via biolistic or 'barley stripe mosaic virus' (BSMV). A report demonstrated that *Pst* effector candidates could be subjected to HIGS in order to reveal their significance in disease formation (Yin et al., 2011). A Mitogen-activated protein kinase (PtMAPK1), a cyclophilin (PtCYC1) and a

calneurin (PtCNB) gene of *P. triticina* (*Pt*) were silenced independently by HIGS methodology in a similar VIGS fashion of dsRNA expression to investigate pathogenicity genes (Panwar et al., 2013b). A recent HIGS study examined a MAPK kinase gene (PsFUZ7) through VIGS mediated delivery and the authors further verified stable expression of siRNA against PsFUZ7 provides effective resistance in wheat (Zhu et al., 2017). Same group reported that a *Pst* gene encoding protein kinase A (PKA) of cAMP-dependent protein kinase (PsCPK1) is also crucial in pathogenesis and stable transgenic wheat lines carrying siRNA for PsCPK1 are able to strongly resist stripe rust disease (Qi et al., 2018).

1.5.5 Heterologous Systems via *Agrobacterium tumefaciens* Mediated Gene Transfer

Investigating effector candidates and characterizing their functions on wheat, (the natural host) is fairly laborious, tricky and challenging due to problematic transient/stable gene expression on host. Biolistic gene delivery is generally in efficient because of thick cuticle layer of wheat. Type 3 secretion system mediated protein delivery still needs optimization and fine-tuning. *Agrobacterium tumefaciens* mediated gene transfer is not possible for wheat in spite of a couple reports stated positive results for delivery of silencing constructs in wheat cultivar, Thatcher (Cuomo et al., 2017; Panwar et al., 2013a).

Model plant systems offer a great chance to scrutinize candidate effectors inside plant cells of well-studied hosts. Chakravarthy et al. described a protocol to screen effector candidates using T3SS and *A. tumefaciens* mediated delivery systems to see if they are able to suppress PTI (Chakravarthy et al., 2009). Strategies relaying on chimeric effector-tagged protein fusions were established to capture interacting partners in vivo utilizing FLAG-tag and fluorescent tags (Petre et al., 2017; Win et al., 2011). Recent publications uncovered the biological roles for numerous effectors of rust fungi fused to fluorescent protein for their interacting partners and subcellular localization inside

model host cells (Lorrain et al., 2018). Petre et al. investigated the subcellular localization and interactome of 20 CSEPs predicted from the genome of the poplar leaf rust fungus *Melampsora larici-populina* (*Mlp*) using *A. tumefaciens* mediated transient gene expression in *N. benthamiana* leaves (Petre et al., 2015). Another 16 *Mlp* CSEPs were predicted and assayed for their localization pattern and their effect on virulence of two pathogen (*P. syringae* DC3000 and *Hyaloperonospora arabidopsidis*) on Arabidopsis thaliana plants by means of stable expression or type 3 secretion system mediated delivery of effectors (Germain et al., 2018). Predicted 16 CSEPs of *Pst* were analyzed through transient expression on *N. benthamiana* leaves for their subcellular localization and their target molecules inside host cells (Petre, Saunders, et al., 2016). A defined pipeline was used to extract 82 CSEPs of Asian soybean rust (*Phakopsora pachyrhizi*) which were further tested in heterologous model systems so as to expose their contribution for suppression of PTI and ETI along with discovery of localization sites (Qi et al., 2017). A paper about haustoria expressed effector candidate (PstHa5a23) demonstrated the competence of the effector in suppressing cell death induced by the transient expression of INF1, BAX1, MKK1 and NPK1 elicitors on *N. benthamiana* leaves (Cheng et al., 2017). In a similar fashion, several cell death inducing R gene/effector complexes were exploited in an attempt to decipher the suppression ability of 20 CSEPs of both *Pst* and *Pgt* as a consequence nine (7 *Pst* and 2 *Pgt*) of the surveyed effectors withheld death occurrence in model system of *N. benthamiana* (Ramachandran et al., 2017). Moreover, one of the effectors Shr7 (suppressors of HR 7) halted the PTI response stimulated by flagellin epitope (flg22) infiltration into *N. benthamiana* leaves (Ramachandran et al., 2017). An effector candidate of *P. pachyrhizi* (PpEC23) suppress HR responses induces by *P. syringae* DC3000 upon delivery with T3SS in multiple hosts (*N. benthamiana, N. tabaccum* cv. Xanthi and *G. max* cv. Williams 82) and additionally PpEC23 was discovered to interact with a transcriptional factor (GmSPL12l) of soybean (Qi et al., 2016).

As we can see, numerous studies have been exploited the advantages of heterologous model systems. Delivery systems for effectors could be transient or stable expression through *A. tumefaciens* or direct protein delivery with T3SS. Assays differ depending on the basis and expectations intended for effector expression.

1.5.6 Subcellular localization studies

The location of a protein is meaningful for evaluation of its biological function. It is expected for a protein to be present in the subcellular location of the interaction site. In similar fashion, pathogen effector needs to travel to the location of its target. Exploiting this phenomenon, the biological function or pathogenicity attribution of an effector could be estimated. For instance, an apoplastic effector is more likely to establish favorable conditions by fighting host defensive measures such as defense enzymes (Section 1.4). If an effector localizes in the nucleolus of host cell, it should be expected to involve in regulation or interference of transcription. Thereby, investigating the localization site of an effector candidate illuminates its role even using model organisms rather than its natural host.

There are reports claiming the effectors manage to mimic host transit peptide sequences for efficient transport inside host cells (Petre, Lorrain, et al., 2016). The presence of transit peptide sequences similar to plant counterparts is intriguing enough to wonder why an effector possess a transit peptide required for chloroplast delivery even though fungus does not have the organelle. Mitochondria and chloroplast are the sites of energy driven processes. However, they also involve in the actions of biotic and abiotic stresses. They produce reactive oxygen species (ROS), phytohormone signals and other defense related activities. Hence, determination of the localization is crucial and informative for effector studies.

Agrobacterium tumefaciens compatible plant destination vectors were designed previously (Karimi et al., 2002). The pK7FWG2 vector (Appendix A) is one of the

destination plasmid having strong 35S promoter site of Cauliflower mosaic virus (CaMV). Cloning the effector of choice into the plasmid result in effector-GFP (Green Fluorescent Protein) fusion at the C-terminal end using flexible 'Gateway cloning' methodology. The versatility of this system was benefitted by numerous researches related to candidate effector investigation (Dagvadorj et al., 2017; Evangelisti et al., 2017; C. Liu et al., 2016; Petre, Lorrain, et al., 2016; Petre, Saunders, et al., 2016).

1.6 Aim of the research

The impact of yellow rust disease on global economy is undeniable due to magnitude of the wheat as a 'Big three crop'. The obligate biotrophic nature of the causative agent *Pst* hampers the efforts to decipher the molecular basis of the disease. Furthermore, working with wheat is also demanding compared to model organisms such as *A. thaliana* and *N. benthamiana* in terms of flexibility in tools of molecular biology. Nevertheless, new advances in both technology and experimentation help us to overcome the obstacles.

In this thesis exertion, the various experimentation strategies are sourced to generate a pioneer transcriptome data of compatible and incompatible interaction occur between *Pst* and wheat. The *Pst* races as a mix collected on wheat leaves in Turkey, which is the homeland for the evolutionary origin of wheat. The efficacy and adeptness of NGS systems are exploited to outline expression patterns by gathering RNA from heavily infected tissues of susceptible and resistant wheat leaves. The time point of 10 dpi (a unique time interval between haustoria formation and sporulation) is chosen to provide distinctive view on disease formation for both compatible and incompatible interaction. Differentially expressed unigenes (DEGs) are hypothesized to be generated by following *de novo* assembly of unigenes subsequent to mapping analysis. A defined pipeline for secretome prediction is constructed to discover the unique sets of small-secreted proteins mined on PstDEGs. The main aim is to discover

a novel catalogue of small-secreted proteins (SSPs) as: 'Differentially expressed small-secreted proteins of *Pst*' or PstDESSPs. Substantial number of tools in effector bioinformatics will be benefitted in a hope to construct an inventory of the most informative and the most promising effector candidates. The data available in literature such as Cantu et al., 2013, Garnica et al., 2013 or Xia et al., 2017, are planned to use as sources for the comparisons against the findings of the thesis. Different comparisons and filtering strategies will assist to determine different sets of suspects. Therefore, the overlapping effectors will be considered as the most promising candidates whereas it does not mean the remaining candidates are less important. A microarray data on RNA material collected from infected leaves of compatible and incompatible interaction will be conducted to provide additional assistance to the expression data of the transcriptome sequencing.

The second aim of the thesis was to determine and establish efficient strategies to assess and analyze highlighted candidates. In order to do so, selected three candidates that were previously studied by Akkaya Research Group previously, are chosen. T3SS delivery systems and heterologous systems to screen effectors contribution on disease in both natural host and model system *N. benthamiana,* are selected to characterize the chosen effectors. The optimization and fine-tuning studies will reveal the fast-forward strategy in terms of time-efficiency and noise. The best strategy will be employed to study the selected three chloroplast or mitochondria candidates from PstDESSPs.

CHAPTER 2

MATERIALS AND METHODS

2.1. Plant and pathogen materials, growth conditions

Wheat cultivars (Avocet-S, Avocet-YR10 and other YR differential lines) were seeded in pots (approximately 300 mL of volume) and grown at 20 ^{o}C of constant temperature along with 16 hours of daytime and 8 hours of nighttime in a climate chamber. *N. benthamiana* seeds were grown in pots sealed with nylon package, which allows breathing and provide extra humidity for better growth. Two weeks-old seedlings were subcultured into bigger pots (300 mL). Conditions for *N. benthamiana* were 24 ^{o}C constant temperature with 60 % relative humidity and 16 hours daytime / 8 hours nighttime.

Figure 2.1. Pst inoculation on two leaf stage wheat leaves. Special apparatus enables spraying urediospore+mineral oil mix on leaf blades (Picture is courtesy of Ayşe Andaç, Akkaya Research Group).

Pst urediospores (fresh) were pre-incubated at 42 ºC for 5 min to increase germination rate. Spores were mixed with mineral oil. With a special spraying tool (Figure 2.1), they were squirted on wheat seedlings at two leaves stage (10-14 days old). Infection was achieved through 10 ºC overnight incubation at dark. Humidifier (steam engines) was used to increase humidity as infection occurs. Following day, the growth conditions were set back to original conditions (20 ºC, 60 % humidity, 16 h daytime). Fresh urediospores were available for harvest after two weeks. Only fresh spores were used for continuous infection upkeep to ensure maximum yield during harvest. Collected spores were treated in a desiccator to stimulate dormant state and then stored in freezer (-80 ºC).

2.2 Tissue sampling and RNA isolation for sequencing

Avocet-S cultivar is susceptible to PstMix whereas Avocet-YR10 cultivar is resistant. We used these wheat cultivars to study compatible and incompatible interactions between *Pst* and the host. Susceptible and resistant cultivars were grown as in Figure 2.2. We had four conditions (3 replicates for each): AvocetS-Mock (ASC), AvocetS-Pst (ASI), AvocetYR10-Mock (AYC) and AvocetYR10-Pst (AYI). The infected leaves were collected using sterile scissors at 10 days post inoculation (dpi) as symptoms were noticeable on the leaves of the susceptible cultivars. The samples were immediately snap-frozen inside liquid nitrogen and stored in freezer (-80 ºC).

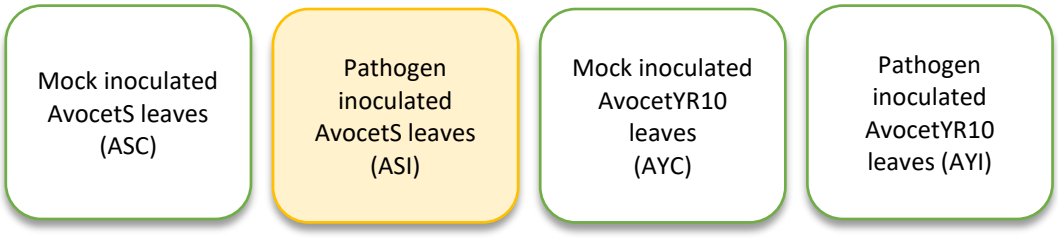


Figure 2.2. The flowchart of the design of the experiments conducted for tissue sampling.

For RNA isolation, 100 mg leaf materials from each replicate of particular condition were crushed together using sterile mortar and pestle. 0.002 g of PVPP (Polyvinylpolypyrrolidone, Sigma CAS no: 25249-54-1) 0.003 g was used for 1 mL of lysis reagent (QIAzol, Cat. No: 79306) during grounding process to ensure high quality RNA yield through removal of polyphenols. The manufacturer's protocol was followed throughout RNA isolations. Briefly, 100 mg of homogenized tissue (mixture of 3 biological replicates) was treated with 1 mL QIAzol lysis reagent for each condition. The samples were never allowed to thaw using liquid nitrogen until addition of lysis reagent. The samples were mixed and incubated at room temperature for 5 min. Subsequently, 0.2 mL of chloroform was added and mixed through inverting. Another incubation was done at room temperature for 2-3 min. The lysates were centrifuged to perform phase separation at 12.000 g for 15 min at 4 ^{o}C. Upper and aqueous phase was transferred to new, sterile tube without disturbing interphase. 0.5 mL of isopropanol was added and mixed through inversions to clean up RNA. Centrifugation was applied at 12.000 g, 4 ^{o}C for 10 min. Supernatant was discarded, and 1 mL of ethanol (75%) was added on the pellet. Centrifugation was performed 7500 g, 4^{o}C for 5 min. Again, supernatant was discarded and the remaining RNA pellet was air-dried. A portion of RNA was dissolved in RNAse free water and the remaining parts were suspended in absolute alcohol for protection during material transfer to sequencing facilities. RNA quality was assessed using spectrophotometer (NanoDrop), gel electrophoresis and Bioanalyzer (Agilent 2100) for observing rRNA bands integrities. RNA integrity number (RIN) in Bioanalyzer analysis was equal or more than 8.0 was the decision rule to proceed.

2.3 Transcriptome sequencing

Sequencing experiments were conducted by Beijing Genome Institute (BGI) as a service. Briefly, samples were subjected to DNAse I treatment, following by mRNA enrichment using oligod(T) magnetic beads. Fragmentation buffer was applied to

obtain short mRNA fragments. These fragments were used as templates for subsequent cDNA synthesis with the aid of random hexamers. Then, cDNAs were purified and dissolved in elution buffer. End reparation and adenine nucleotide addition were performed. Afterward, specific adapters were ligated to short cDNA fragments. PCR amplification was conducted using adapter sequences. Illumina HiSeq™ 2000 sequencing platform was used for deciphering the sequences of constructed library (Protocol was obtained from BGI sequencing report).

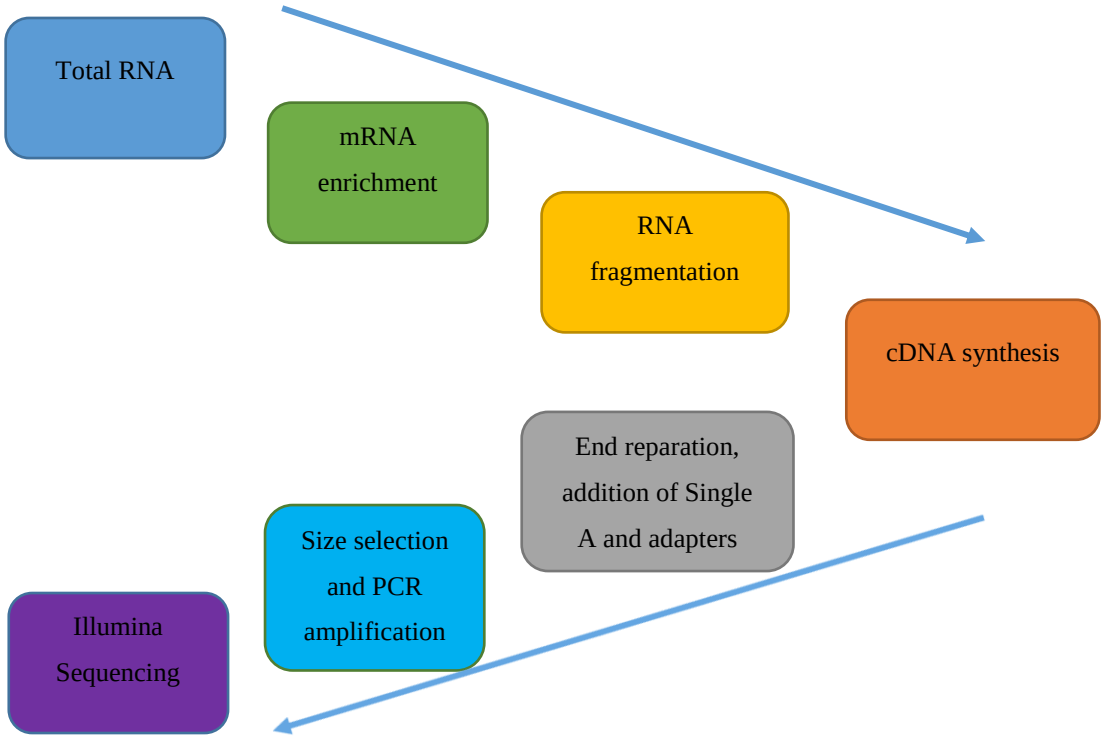


Figure 2.3. Flowchart showing the preparation of total RNA samples for transcriptome sequencing.

2.4 *De novo* transcriptome assembly

Raw reads generated through transcriptome sequencing were subjected to series of bioinformatics analysis. Cleans reads were obtained using following rules: 1) removal of reads with adapter sequences, 2) dismissal of the reads containing unknown/unidentified nucleotide sequences of more than %5 percent of the total sequence, 3) identification and elimination of the low-quality reads that possess more than 20 % reads with the value of $Q \leq 10$. The clean reads were proceeded to assembly stage using Trinity software (release-20130225, http://trinityrnaseq.sourceforge.net/) that uses three different algorithms (Inchworm, Chrysalis, Butterfly) for assembly (Grabherr et al., 2011). Inchworm can assemble the RNAseq data into contigs of unique transcripts. Chrysalis uses Inchworm results to cluster and construct de Bruijn graphs of contigs. Butterfly utilizes Chrysalis graphs of each cluster to comprehend and establish transcript sequences from them in parallel (alternative splicing and paralogous transcripts). Hence, we obtained unigenes as the results of Trinity *de novo* assembly. Based on gene family clustering, unigenes were divided into two subsets: 1) Clusters (group of unigenes that have similarity more than 70 %) and 2) Singletons. The directionality of the unigene sequences were determined by mapping by similarity alignment (Blastx, e-value< 0.00001) against public protein databases of NR (NCBI, non-redundant protein), Swiss-Prot, KEGG and COG (databases are written in priority order of the decision rule if conflict arises between two database). If the direction of a unigene could not be determined by any database, ESTScan (Iseli et al., 1999) program was executed to find out the direction.

2.5 Transcriptome annotation

Aforementioned databases along with some additional ones were used to annotate function and gene ontology classification of the unigenes. Blastx (e-value<0.00001)

was used to search against protein-oriented databases (NR, Swiss-Prot, KEGG and COG) whereas Blastn (e-value<0.00001) for nucleotide oriented (NT) databases (Table 2.1). KEGG database was used to generate pathway annotations (metabolic pathway, cellular process etc.) for gene product predictions of unigenes with internal 'Path_finder' software using default parameters (http://www.genome.jp/) (Kanehisa et al., 2008). Gene ontology (GO) classifications of the unigenes were attributed through Blast2GO program (v2.5.0) using GO database (Conesa et al., 2005). WEGO tool was used to construct GO classifications and annotations to comprehend the distribution of unigenes (Ye et al., 2006). All annotations (Initial annotations) studies described here were done at BGI as a service (Table 2.1).

Table 2.1. The databases used to annotate Unigenes.

Database	**Release**	**Programs and Parameters**	**Reference**
Non-Reduntant Protein (NR) Database	20130408	BlastX (E-value<0.00001)	(Pruitt et al., 2007)
Kyoto Encyclopedia of Genes and Genomes (KEGG)	63.0		(Kanehisa & Goto, 2000)
Clusters of Orthologous Groups (COG)	2013_03		(Tatusov et al., 2003)
Swiss-Prot (EMBL protein database)	20090331		(Bairoch & Apweiler, 2000)
NCBI Nucleotide(NT) Database	20130408	Blastn (E-value<0.00001)	(Pruitt et al., 2007)

2.6 Analysis of expression difference on transcriptome data

Expression difference among compatible and incompatible interaction was analyzed using FPKM (Fragments Per Kilobase per Million reads) method (Mortazavi et al., 2008), which is also equivalent to RPKM (Reads Per Kilobase Million) except for calculation of C and N values. We followed the formula stated below to compute expression levels for each conundrum. The formula is;

$$FPKM = \frac{10^6 C}{NL/10^3}$$

whereas FPKM stands for the number of 'Fragments Per Kilobase per Million reads' generated on the expression level of particular Unigene A (A for any individual Unigene). C signifies the number of fragments that distinctively aligned to Unigene A. N is the total number of fragments that uniquely aligned to all Unigenes and L is the base number in the CDS of Unigene A. As FPKM values for each Unigene was annotated separately, FPKM ratios were calculated between two samples at a time. Two parameters were defined to detect differentially expressed genes (DEGs). First, the ratio should be equal or more than two fold (ratio≥2). Second, false discovery rate (FDR) should be equal or lower than 0.001 (FDR ≤ 0.001). Briefly, 'The significance of digital gene expression profiles' is a strategy to identify DEGs between two samples (Audic & Claverie, 1997). In parallel, a rigorous algorithm was developed to calculate DEGs and filter difference ratio ≥2. FDR is a statistical method to control and examine reliability of p-value in multiple hypothesis testing (Benjamini & Yekutieli, 2001). In practical terms, the FDR is the expectation of getting false a result among total discoveries. For instance, FDR ≤ 0.001 for our study means one or less false discovery amongst thousand discoveries.

Total 10550 unigenes were discovered to fulfill the defined criteria above. These DEGs were further analyze to determine GO terms (p-value ≤ 0.05) including GO functional enrichment and functional classification. Moreover, KEGG pathway

analysis (Q-value ≤ 0.05) was excuted for pathway enrichment to understand biological functions further (Kanehisa et al., 2008).

It should be noted that these DEGs discovered in the transcriptome data of compatible and incompatible interactions belongs to both pathogen and host genes due to de novo assembly. Hence, we will proceed to map *Pst* related and novel genes whereas host related genes will not be covered in this thesis work.

2.7 Mapping of *Pst* related genes and novel genes

The 10550 DEGs discovered through expression difference analysis possesses expression data of both the pathogen and the host. We constructed local database for reference genomes of host and pathogen to map the discovered DEGs. *Pst*-78 genome was used as reference genome for pathogen genes. To minimize the number of unmapped DEGs, we also added the reference genomes of *Pgt* and *Ptt* (evolutionary closest relatives of yellow rust pathogen) and named the database as Pucciniales. Wheat D genome was used to map the host genes.

In summary, whole proteome data of *Pst*-78 (2K41), *Pgt* (strain CRL 75-36-700-3) (race SCCL) and *Ptt* (isolate 1-1, race 1-bbbd) were downloaded from Broad Institute Puccinia website (http://www.broadinstitute.org/) (Cuomo et al., 2017; Duplessis et al., 2011). Wheat D proteome (wheatD_final_43150.gff.cds) of *Aegilops tauschii* was downloaded from GIGA_DB database was download from (http://gigadb.org/) (Jia et al., 2013) and combined with Pucciniales database. The reason for combination is to detect the DEGs that successfully aligned to both Pucciniales and wheat. A local BlastX analysis (Altschul et al., 1997) was conducted via Bioedit program (Hall, 1999) using following parameters: e-value cut-off ≤ e^{-10}. The Unigenes were assigned based on best hits. For instance, a particular Unigene successfully aligned to both Pucciniales and wheat, will be dispersed to the group has best score (decision rule is winner takes all). We categorized DEGs as i) pathogen related, PstDEGs (Unigenes shows

similarity to only Pucciniale or alignment scores are better on Pucciniale compared to wheat), ii) wheat related; TaDEGs (similarity only with or better with wheat) and iii) novel DEGs (Unigenes shows homology to neither on Pucciniales nor wheat). A visual summary of workflow was presented as Figure 2.2.

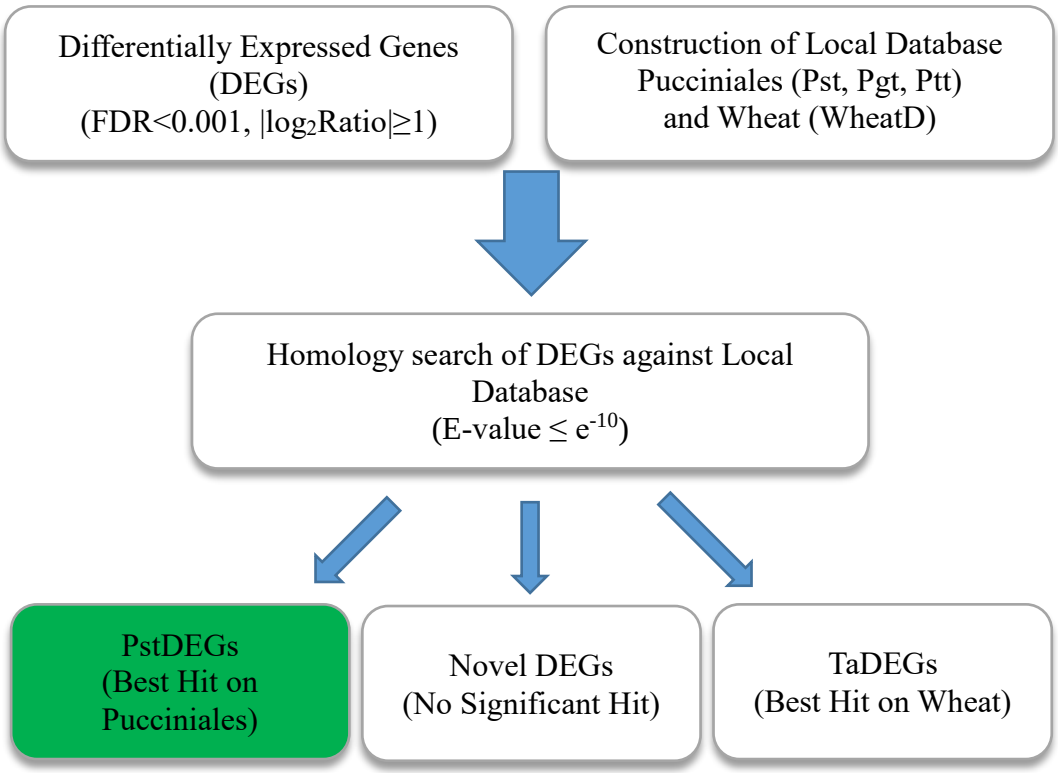


Figure 2.4. Flowchart describes the strategy to identify PstDEGs.

PstDEGs were pioneer data of PstTR-Mix races showing expression levels of the *Pst* genes that is not previously reported in this specific time point (10 dpi) to the best of our knowledge. Therefore, PstDEGs were annotated based on their molecular function, biological process, cellular component, enzymatic activity etc. using Blast2GO program (v 5.1.0).

2.8 Construction of differentially expressed *Pst*-secretome

PstDEGs data were further processed in order to predict secretome repertoire among differentially expressed genes at 10 dpi. A well-studied pipeline (Duplessis et al., 2011; Hacquard et al., 2012; Saunders et al., 2012) with slight changes were applied to filter small secreted proteins (SSPs) from PstDEGs. The following parameters were used as decision criteria: i) presence of signal peptide (SP) sequence at the N-terminus of protein sequence, ii) absence of any transmembrane domain in the mature protein (without SP), iii) mature protein length should be equal or smaller than 300 amino acids.

The open reading frame (ORF) predictions were achieved using ORFPredictor webtool (http://bioinformatics.ysu.edu/tools/OrfPredictor.html). For signal peptide prediction, Signal P (v 4.1) web tool (Petersen et al., 2011) was used using default parameters (http://www.cbs.dtu.dk/services/SignalP/). Transmembrane domain predictions were performed using TMHMM web tool (v 2.0) (http://www.cbs.dtu.dk/services/TMHMM/) (Krogh et al., 2001; Möller et al., 2001; Sonnhammer et al., 1998)(Moller et al., 2001). The consequential list was named as PstDESSPs (Differentially Expressed Small Secreted Proteins of *Pst*, at 10 dpi).

2.9 *In silico* characterization of PstDESSPs

The characterization was achieved using *in silico* analysis to forecast wide range of attributes including function, subcellular localization, domains, motifs, effector predictions etc. The details of the analysis are listed in each section below and described briefly.

2.9.1 Function annotations

The function annotations were conducted using Blast2GO program or BlastKOala webtool to annotate functions. However, the enzymatic activities such as proteases and protease inhibitors, lipases, CAZymes or oxidoreductases were further predicted by similarity search against public databases constructed specially on these enzyme classes. The protein sequences of the PstDESSPs were analyzed for their similarity to known protease and protease inhibitors via alignment search (E-value<$1.0e^{-5}$) on MEROPS database (http://merops.sanger.ac.uk/) (Rawlings et al., 2018). CAZymes were annotated based on homology search on dbCAN database (http://csbl.bmb.uga.edu/dbCAN/) (Yin et al., 2012). Lipase activities were predicted via homology search on LED database (http://www.led.uni-stuttgart.de/). Oxidoreductases were downloaded from 'fPoxDB' database that constructed on fungal peroxidases as a platform (http://peroxidase.riceblast.snu.ac.kr/) (Choi et al., 2014). BlastP search (E-value<$1.0e^{-5}$) was conducted on downloaded fungal peroxidase to annotate PstDESSPs using BioEdit program. The results were combined as Table in Appendix C using Microsoft Excel 2016.

2.9.2 Subcellular localization predictions

Subcellular localization predictions were conducted on mature proteins under the assumption that effector proteins must lose its signal peptide sequence upon secretion and translocate into host without SP. Therefore, mature protein sequences of PstDESSPs were analyzed for their subcellular localization using TargetP, Localizer, WolfPSORT and ApoplastP web tools. For TargetP (v 1.1) predictions were achieved using default parameters for plant option with 'winner takes all' decision rule (http://www.cbs.dtu.dk/services/TargetP/) (Emanuelsson et al., 2000). For Localizer (v 1.0.4) predictions, 'mature proteins without signal peptides' option was chosen and default parameters were used (http://localizer.csiro.au/) (Sperschneider et al., 2017).

WolfPSORT was executed by choosing ‘plant’ option with default parameters (https://wolfpsort.hgc.jp/) (Horton et al., 2007). Likewise, ApoplastP (v 1.0) was operated using default parameters (http://apoplastp.csiro.au/) (Sperschneider et al., 2018). All predictions were added to Table (Appendix D).

2.9.3 Conserved domain analysis

Conserved domain analysis was performed using ‘Conserved Domains Database’ (CDD) of ‘The National Center for Biotechnology Information’ (NCBI) which amasses the data of several external sources (Pfam, SMART, COG, PRK, TIGRFAMs) along with NCBI curated protein data (Marchler-Bauer et al., 2011; 2015 and 2017). The analysis was executed on mature proteins with default parameters of all results modes using CD-Search web tool (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (Marchler-Bauer & Bryant, 2004). The results were downloaded in concise, standard and full format. Only the concise results were processed as conserved domain due to large number of domain hits obtained for single protein. Concise results option eliminated all other hits that survive the expect value threshold. These results were merged to Table in Appendix C.

2.9.4 Virulence analysis

The virulence or pathogenicity attributes of PstDESSPs were scanned through ‘pathogen host interaction database’ (PHI-base) which was available as a web tool using BLAST strategy (PHIB-Blast) to characterize candidate effectors (http://phi-blast.phi-base.org/) (Urban et al., 2015). The PstDESSPs were scanned for their similarities to the data (v 4.1, 4.2, 4.3, 4.4 and 4.5) of PHI-base using default

paramaters. The collected results were further processed using cut-off or expect value (E-value<e^{-5}) and marked green in Table (Appendix E).

Another fungal virulence factor database (Lu et al., 2012) were downloaded from the website (http://sysbio.unl.edu/DFVF/) and local BlastP analysis (E-value<e^{-5}) was conducted on PstDESSPs using BioEdit program. However, the data was not included since low percentage of positive hits. In addition, database is not updated since its release.

2.9.5 Effector predictions

The deep machine-learning algorithm based EffectorP web tool (http://effectorp.csiro.au/) was used to predict effectors in PstDESSPs. Initial analysis was done using EffectorP (v 1.0) with applying default parameters (Sperschneider et al., 2016). However, updated version of the EffectorP (v 2.0) algorithm was offered during thesis work (Sperschneider et al., 2018). Similar pattern was applied using default parameters. The results were organized and merged in Table (Appendix E) and marked as orange for positive and magenta for 'unlikely effector'.

Due to cysteine coverage of candidate effector matters, we manually counted Cysteine residues using 'Notepad ++' program. The decision rule was cysteine percentage should above 3% of the total amino acid number. The results were annotated in Table (Appendix C).

2.9.6 Comparisons with published works

'Haustoria expressed secreted proteins' (HESPs) and 'infected leaves expressed secreted proteins' (ISEPs) were downloaded as tribe sequences from supplementary materials of Cantu *et al.*, 2013. Local BlastP analysis was conducted using BioEdit program with a described cut off value (E<e-^{10}) for similarity search. The results were processed using internal Perl scripts in Notepad ++ and listed in Appendix F.

The avirulence (Avr) candidates generated by genomic correlation analysis were obtained from Xia et al., 2017. The Avr candidates were listed as *Pst*-78 proteins. Matching Pst-78 homologs of PstDESSPs were listed in Appendix F and listed as Avr candidates for YR lines.

'Species specific secreted proteins' (SSSPs) that were data mined through kingdom-wide analysis, were downloaded from Kim et al., 2016. BlastP similarity search was performed using BioEdit program ($E<e^{-10}$) on PstSSSPs described by Kim et al., 2016. The results were included in Appendix F.

The transcripts data of germinated urediospores and haustoria that published by Garnica et al., 2013, were downloaded from supplementary materials of the paper. Local BlastX analysis ($E<e^{-10}$) were conducted on PstDESSPs with BioEdit program. The results were included in Appendix F.

Among PstDESSPs, the subset of EffectorP positive, cysteine-rich, homologs of HSEPs and ISEPs (Cantu et al., 2013) were compared to observe the most promising candidates. The venn diagram was drawn using by a web tool (http://bioinformatics.psb.ugent.be/webtools/Venn/) to visualize overlapping subset of PstDESSPs.

2.9.7 Phylogenetic tree analysis

Evolutionary relationship and conserved patterns were analyzed using multiple sequence alignments strategy. The ClustalW program was operated to analyze homology amount using default parameters (https://www.ebi.ac.uk/Tools/msa/clustalo/) (Sievers et al., 2011). The results were downloaded in Newick format. The visualization through phylogenetic tree was achieved using iTOL web tool (https://itol.embl.de/) (Letunic & Bork, 2016). Newick format of multiple sequence alignment was uploaded to iTOL. Circular phylogenetic

tree was constructed. Observed three main branches were colored differently to ease further analysis.

2.9.8 Motif discoveries

The conserved motifs among three main branches discovered through multiple sequence alignment, were analyzed using MEME web tool (http://meme-suite.org/tools/meme) (Bailey et al., 2009). The parameters were set as '10 motifs', 'zero or one occurrence per sequence', 'minimum width is 3, minimum sites were chosen 50 and 100' and the remaining parameters were set as the default. Another search was repeated as 'minimum width is 6, minimum sites were chosen as 20 and 50'. The results were recorded and downloaded.

The previously known/reported motifs ([F/Y/W]xC, RxLR etc) were searched manually using 'Notepad++' program. The results were integrated to the table in Appendix C for better comparison and overall look.

2.10 RNA isolation for microarray assays

RNA isolation for microarray analysis was conducted following the protocol aforementioned in Section 2.2. The pathogen race used in the study was PstTR0997 (An isolate in PstTRmix which is also virulent to AvocetS whereas avirulent on AvocetYR10). The growth conditions and inoculation protocol were same as the described protocol in Section 2.1 and 2.2. The samples were collected at 24 hpi (hours post inoculation), 72 hpi and 10 dpi. Three biological replicates for each time interval and disease condition were harvested as in Figure 2.2. The samples were snap-frozen in liquid nitrogen and stored at deep freezer (-80^{o}C) until RNA isolation. Total RNAs were isolated from six condition (24 hpi ASC, 24 hpi ASI, 10 dpi ASC, 10 dpi ASI, 10 dpi AYC and 10 dpi AYI) using the QIAzol protocol described in Section 2.2

except none of the resulting RNA was dissolved in absolute ethanol due to no material transfer was required this time. In addition, three replicates for each condition were homogenized together as in Section 2.2 and divided into three technical replicates respectively. Total RNA was treated with DNAse I to remove residual genomic DNA by following the manufacturer's protocol. The quality of total RNA for each sample was assessed by following methods: visualizing the integrity of rRNA bands on agarose gel electrophoresis, quality and quantity assessment on micro-spectrophotometer (NanoDrop) and BioAnalyzer tool (RIN>6).

2.11 Microarray analysis

Microarray probes for pathogen were designed based on Pst-78 genome. Host gene expressions were analyzed using probes based on *T. aestivum*. Sixteen chips were produced using customized probes by Agilent Company as service. Probes on chips were designed in 8x60k format by 'Agilent eArray design tool' following the base composition methodology. The probes design options were settled using best distribution methodology. Three chips were used for 10 dpi samples of ASC, ASI, AYC and AYI. Two chips for 24 hpi samples of ASC and ASI. The experiments were conducted as service by Agilent.

The results were subjected to normalization and statistics analysis for assessment and comparison. Briefly, BRB-ArrayTools (v 4.3.2) was the program of choice in order to analyze the data collected through microarray (Simon et al., 2007). BRB-Array tool was executed using R software with Bioconducter package as a Microsoft Excel Plug-in. The results were displayed as a table of Microsoft Excel program. Briefly, initial data of the microarray study were normalized by means of quantile normalization to remove background noise arisen out of technical glitches. Boxplots were constructed to visualize the differences before and after the normalization process. Associations between pathogen challenged plants and their controls (mock treatment) were analyzed with 'Class comparison test' pursuing the decision rule ($p \leq 0.05$ and 2-fold

change). Univariate test of two-sample t-test with random variance model was executed as class comparison command. The results were sorted as differentially expressed genes among classes in Excel files. The significant genes showed at least two fold change, were visualized using volcano plots. Cluster analysis was conducted on the significant genes (10-fold change) using hierarchical 'Average linkage' method wih Cluster (v 3.0) program (http://bonsai.ims.u-tokyo.ac.jp/~mdehoon/software/cluster/) (Eisen et al., 1998). The results of cluster analysis were visualized using Java TreeView (v 1.1.6r4) program.

Due to recent progresses in *Pst*-78 reference genome, we re-aligned probes to *Pst*-78 predicted transcripts to obtain updated list of pathogen genes. The results of microarray analysis were compared with *de novo* transcriptome results (PstDEGs) with local Blastn analysis of BioEdit program. However, too few of them matched with PstDEGs. Therefore, we separated results of microarray analysis and evaluated them independently.

2.12 Pilot studies for candidate effector characterizations *in vivo*

Transcriptome analysis and microarray studies were proceeding, we started to work with previously studied effector candidates in our research group (Ozketen A. C. 2013, Master of Science thesis) (Andac A. 2013, Master of Science thesis). Pstha15N21, Pstha12H2 and Pstha12J12 were previously cloned into pEDV6 vector in *P. fluorescens* EtHAn. In this section, we described the protocols to study these candidate effectors with *P. fluorescens* EtHAn mediated wheat infiltration assay (WIA).

2.12.1 *P. fluorescens* EtHAn mediated wheat infiltration assays

The protocols for *P. fluorescens* EtHAn mediated WIA experiments were obtained from Dr. N. Upadhyaya (adapted version from David Joly, Canada) (Upadhyaya et al., 2014). Briefly, starter cultures were initiated with inoculations of *P. fluorescens*

EtHAn (containing pEDV6-PstHa15N21, pEDV6-PstHa12H2, pEDV6-PstHa12J12 and mock bacteria) clones from stock culture into 5 mL LB (Lysogeny broth) medium (See Table 2.2) with Gentamycin (100 µg/mL)/Chloramphenicol for empty bacteria. The cultures were incubated for overnight at 29 ºC at 200 rpm. The grown cells were sub-cultured into new LB medium of 50 mL volume with Gentamycin (100 µg/mL) or Chloramphenicol for 4-6 hours at 200 rpm until cell density reaches A_{600nm} value between 1.3-1.6 for optical density.

Table 2.2. Ingredients for LB and LB agar preparations.

Reagents	**Amounts**
Peptone	5 g
Yeast Extract	2.5 g
NaCl	5 g
NaOH (0.5 M)	1.6 mL
Agar (for LB Agar)	7.5 g
ddH_2O	Upto 0.5 L
Sterilize the mixture by autoclave.	

The bacteria were centrifuged at 5000 g for 7 min. at 4ºC. The medium (supernatant) was discarded. The harvested cells were washed with 25 mL of 10 mM $MgSO_4$ (ice-cold) twice by repeating centrifugation and washing steps. Last, the cells were resuspended in minimal media (See Table 2.3) and cell density was adjusted at 1.0 at A_{600nm}. Overnight incubation was performed at 20ºC, 200 rpm. Next day, the cells were collected by centrifugation at 5000 g, 4ºC for 7 min. Finally, the bacteria cells were dissolved in 10 mM $MgCl_2$ solution. Resuspended cells were adjusted to 2.0 of optical density at A_{600nm} for final tuning.

Table 2.3. Recipe for minimal media preparation.

Reagents	**Initial Concentration**	**Volumes**	**Final Concentration**
K2HPO4 (Merck, Lot # A488700-023)	1 M	4.25 mL	50 mM
KH2PO4 (Merck, Lot # A0104673-938)	1 M	45.75 mL	50 mM
(NH4)2SO4 (Merck, Lot # A897317-824)	760 mM	10 mL	7.6 mM
MgCl2 (Sigma, Lot # 49H04351)	1.7 M	1 mL	1.7 mM
NaCl	1.7 M	1 mL	1.7 mM
Sucrose (Merck, Lot # K41912651-128)	Upto 0.5 L	3.42 g	10 mM
Sterile H2O		Upto 1 L	-
Adjust pH (5.7-5.8) Sterilize the medium by filter (0.20 µm) sterilization			

The yellow rust (YR) differential lines used for WIA were listed in Table 2.4. The YR differential lines were grown until day reach 2-3 leaves stage in specially designed seedling pots following the protocol described in Section 2.1. The each line was infiltrated at the bottom of the second leaf blade with *P. fluorescens* EtHAn clones with approximately 0.2 mL volume with the aid of needleless syringe. The infiltrated areas were marked and let dry for a brief period. The infiltrated differential lines were placed in a growth chamber (Panasonic) and WIA program (60 % humidity, 24 ºC dark) was executed with three vials of hot water to provide misty environment for 1 hour. Next day, normal growth conditions (60 % humidity, 16 h daytime, 20 ºC) for wheat was applied and hypersensitive responses were detected on second day and third day. The results were photographed and recorded.

Table 2.4. YR differential lines

YR Differential Lines	Known R genes
Siete Cerros T66	YR2
Avocet-YR18	YR18
Avocet-YRSP	YRSP
Pavon F76	YR29,YR30+1
Seri M 82	YR2, YR9, YR29, YR30, +
Super Kauz	YR9, YR27, YR18
Polmer 2.1.1	?
Avocet-YRCV	YRCV
Avocet-YR*3/Altar 84/AE.SQ/OP	YR27,YR18
Avocet-YR*3/LALMONO1/PVN	?
Avocet-YR*3/Pastor	?
Jupateco R	YR18
Kalyansona	YR2

WIA on YR differential lines of Avocet near isogenic lines

The WIA assay described above was repeated on new sets of differential lines. The YR lines were produced on AvocetS cultivar and they differ only in YR resistance. The complete list is in Table 2.5.

Table 2.5. YR differential lines of Avocet near isogenic lines.

YR Lines
Siete Cerros T66
YR1/6*AvocetS
YR5/6*AvocetS
YR6/6*AvocetS
YR7/6*AvocetS
YR8/6*AvocetS
YR9/6*AvocetS
YR10/6*AvocetS
YR11/3*AvocetS
YR12/3*AvocetS
YR15/6*AvocetS
YR17/6*AvocetS

2.12.2 Optimizations of WIA assays

False positive results were revealed in second trial of WIA. The optimization experiments were conducted to solve the inconsistency among the results of the WIA trials in Section 2.11.1. The best cell density was investigated to stimulate HR on effector candidates while not awakening any HR on control groups. The WIA protocol reported in Liu et al., 2016 was performed. The step that includes minimal media induction was omitted from the procedure. *P. fluorescens* cells were grown on KB agar with corresponding antibiotics and directly scraped from agar media with the aid of sterile pipette tips. The final concentration of the cells before infiltration was adjusted between 1.0 and 2.0 at A_{600nm} by suspending in sterile $MgCl_2$ (10 mM) solution. The Siete Cerros T66 cultivar was chosen for infiltration due to positive results observed in first trial. The results were recorded at 4^{th} day after infiltration.

2.12.3 Cloning of eGFP into pEDV6 vector

The cloning of eGFP was achieved through two steps. First, eGFP gene sequence was amplified from pK7FWG2 vector using appropriate primer sequences. Second, it was cloned into pEDV6 vector by 'Gateway cloning method'.

Amplification of eGFP

eGFP primers were designed as CACC sequence at the 5'-end of the forward primer and stop codon TGA at the 3'-end of the reverse primer (See Table 2.12). The designed primers were inspected for their 5' and 3' end homodimer and heterodimer formation with 'oligoanalyzer' tool, which is accessible online in the IDT DNA website (https://sg.idtdna.com/site/account/login?returnurl=/calc/analyzer), and melting temperature values were checked in 'NEB (New England Biolabs Inc.) Tm calculator' online tool publicly available (https://tmcalculator.neb.com/). The PCR amplification of eGFP sequence from pK7FWG2 vector was achieved by Q5 polymerase enzyme (NEB, Cat #M0478G) reaction by following the manufacturer's protocol described in Table 2.6 using thermal cycler (Eppendorf) conditions listed in Table 2.7.

Table 2.6. Polymerase Chain Reaction (PCR) used for amplification of eGFP.

Reaction Component	**Volumes**	**Final Concentration**
5X Q5 Reaction Buffer	5 µl	1X
10 mM dNTPs	0.5 µl	200 µM
10 µM Forward Primer	1.25 µl	0.5 µM
10 µM Reverse Primer	1.25 µl	0.5 µM
Template DNA (Nuclease free water for the negative control)	1 µl	<1000 ng
Q5 HF DNA Polymerase	0.25 µl	0.02 U/µl
5X Q5 High GC Enhancer	5 µl	1X
Nuclease-free Water	Upto 25 µl	

Table 2.7. Reaction Parameters for Thermal Cycler.

Reaction Steps	**Temperature**	**Time**
Initial Denaturation	98 ºC	2 minutes
35 Cycles (Denaturation, annealing and extension)	98 ºC	10 seconds
	55 ºC	30 seconds
	72 ºC	30 seconds
Final Extension	72 ºC	2 minutes
Hold	4 ºC	

The reaction products were separated using agarose gel (1 % Agarose in TAE buffer with EtBr) electrophoresis (60-70 V for 50-60 min). The DNA bands were visualized under UV light (Kodak, Gel Image System) and recorded.

pENTRY/D-TOPO cloning

The amplificant of PCR was cloned into pENTRY/D-TOPO vector by following the protocol of the manufacturer. Briefly, the reaction mixture was prepared as in Table 2.8. The pENTRY/D-TOPO plasmid was added gently without pipetting or disturbing viciously due to topoisomerase is attached to the plasmid itself with covalent bond. The reaction was incubated at 22-23 ºC for 30 minutes. Then, the mixture was incubated at 4 ºC for overnight.

Table 2.8. The components of pENTR/D-TOPO cloning reaction.

Reaction Components	**Volumes**
PCR product	1 µL
pENTR/D-TOPO	1 µL
Salt solution	0.5 µL
Nuclease-free water	3.5 µL

Preparation of heat shock competent cells using *E. coli* Top10

A single colony of *E.coli* TOP10 (Invitrogen) was cultivated using overnight incubation at 37 ^{o}C, on LB agar plate by streak plate technique. The single colony was transferred to 4 mL of LB medium by sterile loop inoculation. The incubation was performed at 37 ^{o}C for overnight (16-18 hours) at 200 rpm. Next day, 0.5 mL of grown culture was transferred in 50 mL of LB medium. The incubation was performed at 37^{o}C, 200 rpm until the optical density value of the cell culture reaches to 0.375 at A_{600nm}. The cells were treated in ice bath to end growth phase. Subsequently, the cells were harvested by centrifugation at 5000 rpm, 4 ^{o}C for 5 minutes. Collected cells were washed twice with 10 mL of $CaCl_2$ (100 mM) solution, which was sterilized with filter (0.2 µm) and pre-chilled on ice. The pellet obtained after washing and centrifugation steps was resuspended in 2 mL of $CaCl_2$ (100 mM) and stored at 4 ^{o}C. The widely used protocol was to store the competent cells of 50-100 µl volumes treated with glycerol in sterile vials. However, we chose to use fresh prepared competent cells, which preserve their competency for 10 days.

Heat shock transformation

Reaction product obtained in pENTR/D-TOPO cloning above was transformed into competent cells of *E. coli* Top10. The product was mixed into 100 µl of *E. coli* Top10 cells by gently pipetting. The cells were placed in ice bath and incubated for 10 min. Then, the vials were swiftly located in water bath (pre-heated to 42 ^{o}C) for 45-60 seconds. The samples were immediately positioned back in ice bath for 5 min to complete heat shock transformation. LB medium (500 µl) was added to the vials. The incubation was applied for 1-2 hours at 37 ^{o}C, 200 rpm. After recovery phase, the cells were centrifuged shortly for 2 minutes. Approximately 400 µl of LB was removed from the mixture and the remaining cells were suspended. This step was optional; it was conducted to concentrate positive transformants on single, selective agar media.

The cells were plated on selective LB agar plate (Kanamycin, 50 µg/mL) and incubated at 37 ºC for overnight.

Colony PCR

Positive results observed on plates were subjected to colony PCR validation to screen false positives. The primers for eGFP were used for colony PCR. The reaction mixture was prepared as in Table 2.9 with negative control. The protocol was obtained from NEB Company's website.

Table 2.9. List of the reactants used in colony PCR.

Reaction Component	**Volumes**	**Final Concentration**
10X Standard *Taq* Reaction Buffer	2.5 µl	1X
10 mM dNTPs	0.5 µl	200 µM
25 mM MgCl2	1.5 µl	0.5 µM
10 µM Forward Primer	0.5 µl	0.5 µM
10 µM Reverse Primer	0.5 µl	
Template DNA (Nuclease free water for the negative control)	1 µl	<1000 ng
Q5 HF DNA Polymerase	0.125 µl	0.02 U/µl
Nuclease-free Water	Upto 25 µl	

Table 2.10. Parameters for the setup of Thermocycler machine.

Reaction Steps	**Temperature**	**Time**
Initial Denaturation	95 ºC	2 minutes
35 Cycles (Denaturation, annealing and extension)	95 ºC	30 seconds
	55 ºC	30 seconds
	68 ºC	40 seconds
Final Extension	68 ºC	5 minutes
Hold	4 ºC	

Plasmid isolation

The positive colony verified with the colony PCR was inoculated in LB medium with kanamycin (50 µg/mL) and incubated at 37 ºC, 200 rpm for overnight. The grown culture was harvested though centrifugation at 5000 rpm using regular benchtop centrifuge. The supernatant (medium) was discarded. The remaining protocol was based on QIAGEN Spin Miniprep kit (Cat # 27106, Lot # 139315519). P1 (RNAse A added) solution (250 µL) was added to the pelleted cells. The mixture was suspended by vortex and shaking. P2 (LyseBlue added) solution (250 µL) was poured into reaction and mixing was achieved by gentle inversions (10-20 times) until blue color was homogenized. N3 buffer (350 µL) was added to the reaction to neutralize. The mixture was shaken through gentle inversions (10-20 times) until blue color disappears completely. The reaction was poured into silica columns supplied by the manufacturer and centrifugation was done at 13000 g. The flow through was thrown away. The PB buffer (500 µL) was applied to the column and centrifuged at 13000 g. The flow through was removed. The PE (ethanol added) buffer (750 µL) was used to wash the column and centrifugation was applied at 13000g. Sterile, nuclease-free water was used to elute the plasmid DNA. The quality of plasmid DNA was assessed at micro spectrophotometer (NanoDrop).

Cloning to Destination Vector

Gateway cloning LR strategy was exploited to subclone ‘gene of interest’ from pENTR/D-TOPO to pEDV6 destination vector. The reaction kit was Gateway LR Clonase II Enzyme Mix (Invitrogen, Ref # 11791-043, Lot #1671682). The protocol was tailed as the manufacturer instructed in Table 2.11. The ratio of the pENTR/D-TOPO to pEDV6 was 1:1 in terms of concentration. The reaction was incubated at 25 ºC for 1 hour and at 4 ºC for overnight.

Table 2.11. The reagents for LR cloning.

Reaction Components	Volumes
pENTR-eGFP	2 µL
pEDV6	2 µL
2X TE Buffer	5 µL
LR clonase	1 µL

Electrocompetent cell preparation of *P.fluorescens* EtHAn

Electrocompetent cell preparation method was adapted and modified from the protocol designed for *P. aeruginosa* (Choi et al., 2006). Briefly, a single cell culture of *P. fluorescens* EtHAn was initiated on KB (Kings Broth) agar (chloramphenicol 35 µg/mL, optional) from master (-80 ºC) stock using streak plate technique for 29 ºC through overnight incubation. KB medium was prepared by mixing 1 g peptone, 1 mL glycerol, 0.15 g K_2HPO_4, 0.5 mL $MgSO_4$ (1M) and 1.5 g agar (only for KB agar preparation) in 100 mL and sterilized with autoclave. The colony was inoculated into 6 mL of KB medium (chloramphenicol 35 µg/mL, optional) and incubated at 29 ºC, 200 rpm for overnight. Following day, the grown cells were separated into four sterile centrifuge tube. The centrifugation was performed at 16000 g at room temperature for 2 min. The supernatant was discarded and the remaining pellet was suspended by pouring 1 mL of sterile sucrose (300 mM) solution. Sterilization was achieved using micro filters (0.20 µm). The centrifugation was applied at 16000 g for 2 min. The washing step was repeated again with 1 mL of sterile sucrose (300 mM) solution. After final centrifugation, the collected cells were resuspended in sucrose solution and combined together in a single sterile vial. The fresh electrocompetent cells were used prior to the transformation with electric shock.

Transformation of pEDV6-eGFP into *P. fluorescens* EtHAn

Transformation was executed following the electroshock transformation method. The freshly prepared electrocompetent cells (100 µL) were mixed with 5 µL (100-200 ng DNA) of pEDV6-eGFP plasmid. The mixture was poured into sterile pulser cuvette (Celljject duo Electroporator, Thermoscientific) which was pre-cooled in refrigerator prior to use. The cells in the cuvette were zapped at 25 µF, 2.2 kV with single electric pulse. The resultant cells were transferred to sterile vial having 1 mL of KB medium and incubated for 2 hours at 200 rpm at 29 ^{o}C. After recovery phase, the cells were centrifuged briefly at 4000 rpm for 2 min. The supernatant was removed until approximately 200 µL of medium remains. The cells were resuspended and spread on selective LB agar with Gentamycin (100 µg/mL). The incubation was performed at 29 ^{o}C for overnight. The positive clones were subjected to the colony PCR verification as described previously.

2.12.4 *P. fluorescens* EtHAn mediated PTI assays

The ability of effector candidates were assayed using various strategies on different hosts (*N. benthamiana* and *Triticum aestivum*). *P. fluorescens* EtHAn and *A. tumefaciens* were used for effector delivery whereas *P. syringae* was used for cell death inducer.

P. fluorescens EtHAn mediated PTI assays on tobacco

The protocol that chosen for PTI suppression assay was obtained from Chakravarty et al., 2009 with slight changes in the procedure. The bacteria strains were plated on their corresponding selective media. *P. fluorescens* EtHAn was spread on KB agar (chloramphenicol 35 µg/mL) and *P. fluorescens* EtHAn carrying pEDV6-PstHa15N21, pEDV6-PstHa12H2, pEDV6-PstHa12J12 and pEDV6-eGFP were grown on KB agar with Gentamycin (100 µg/mL) at 29 ^{o}C. *P. syringae* DC3000 strain

was (obtained from Dr. Joe Win) incubated on KB agar with Rifampicin (25 µg/mL) at 29 ^{o}C. The grown cells of *P. fluorescens* were collected from KB agar with the side surface of a sterile pipette tip in sterile water. The bacteria were washed and centrifuged twice in $MgCl_2$ (10 mM). The optical density (A_{600nm}) was adjusted to 0.5 for *P. fluorescens* EtHAn and 0.02 for *P. syringae* DC3000. The upper leaves of *N. benthamiana* (4-6 weeks old) were chosen for assay and infiltrated with *P. fluorescens* EtHAn (empty or carrying a vector) through the bottom side of the leaf blade with needless syringe. The infiltration area was marked as circles with a permanent marker and plants were kept at room temperature for drying. After seven hours, *P. syringae* DC3000 was challenged in similar fashion described here. The infiltration was performed by carefully overlapping the challenge circle and previous circle. The sides of the overlapping area should be at the center of first and second inoculations. The plants were incubated in growth chamber by following the routine parameters described in section 2.1. The results were recorded at 2-3 days post inoculations.

P. fluorescens EtHAn mediated PTI assays on wheat

Two protocols for the assay that reported by two different groups, were combined and adapted (Lie et al., 2016) (Ramachandran et al., 2016). The *P. fluorescens* EtHAn (chloramphenicol 35 µg/mL) and *P. fluorescens* EtHAn carrying pEDV6-PstHa15N21, pEDV6-PstHa12H2, pEDV6-PstHa12J12 and pEDV6-eGFP were grown on KB agar with Gentamycin (100 µg/mL) at 29 ^{o}C for overnight. The grown cells were collected with a sterile pipette tips as mentioned in the previous part. The bacteria were washed by suspending in filter-sterilized MgCl2 (10 mM) solution followed by centrifugation, twice. The final cell concentrations for *P. fluorescens EtHAn* were attuned to 1.0 and for *P. syringae* DC3000 to 0.3 at A_{600nm}. Each one of all *P. fluorescens* EtHAn clones (control, pEDV6-PstHa15N21, pEDV6-PstHa12H2, pEDV6-PstHa12J12 and pEDV6-eGFP) were combined with *P. syringae* DC3000 in 1:1 volumetric ratio in advance of the infiltration process. The inoculations were

achieved with sterile, needless syringe on the bottom side of the leaf blades of wheat (10-14 days old, two-leaf stage) lines (Avocet-S, Kalyansona-Yr2 and Avocet-YR10). The infiltration areas were marked with permanent marker. The plants were placed in a growth chamber at regular growth conditions stated in section 2.1. The results were collected and recorded at 3rd days post inoculation.

A. tumefaciens mediated PTI assays on tobacco

A. tumefaciens GV3101 clones (carrying pJL48-PstHa15N21, pJL48-PstHa12H2, pJL48-12J12 and pJL48-GFP) were constructed and provided by 'Akkaya Research Group'. The bacteria were plated on selective LB agar plate with Kanamycin (50 µg/mL) and incubated at 28 ^{o}C for 2 days. The grown cells were scratched from the surface of the LB agar by the aid of a sterile pipette tip. The bacteria were collected in sterile water and washed twice. Centrifugation was applied at 4000 rpm for 2 min. An additional washing step was applied using 'Agroinduction media' described in Table.

2.13 Amplification and cloning of candidate effectors

The effector candidates from Appendix B were chosen randomly. The primer designs were done based on constructs will be cloned in Gateway vector with or without signal peptide region in their N-terminus end and without a stop codon in their C-terminus region to allow GFP fusions. The primer list were used in the study was presented in Table 2.12.

Table 2.12. Primers used in this study

Names of The Primers	**Sequences (5'-3' direction)**
CACC-SP-730F	CACCATGTTCCTCGTCTTGACGTTT
CACC-730F	CACCATGTCTTCAATCCAAATGTGTG
730Rev-STP	TTAATTTGTGACAGGGTCACAGTTG
730Rev	ATTTGTGACAGGGTCACAGTTGG
CACC-SP-917F	CACCATGTTGTTCTACGTTTACCTCA
CACC-917F	CACCATGCAGACTTTACCTTCCG
917Rev-STP	CTAGCATGTTTCCCAGCCTCC
917Rev	GCATGTTTCCCAGCCTCCG
CACC-SP-651F	CACCATGACATTAGGGACTTTGACT
CACC-651F	CACCATGTCGCCGCTACCCTC
651Rev-STP	TTAACTCACAGGTAGGGTTCCTGT
651Rev	ACTCACAGGTAGGGTTCCTGTTTTG

2.13.1 cDNA synthesis

RNA isolation was performed in section 2.2. The stored RNA samples (24 hpi, 72 hpi and 10 dpi ASI samples) were used as template for cDNA synthesis. Prior to use, DNase I treatment was performed to ensure the removal of residual genomic DNA. One µg RNA was treated with the enzyme and the reagents listed in Table. Incubation was done at 37 ºC for 30 min. The enzyme was inactivated by adding 1 µL of EDTA (50 mM) and incubated at 65 ºC for 10 min.

Table 2.13. The ingredients for DNase I treatment.

Reaction Components	**Volumes**
Total RNA	1 µg
10X Reaction Buffer (with MgCl2)	1 µL
RNAse-free DNase I	1 µL (1U)
Nuclease-free water	Upto 10 µL

The reaction products were used as template for first strand DNA synthesis. The primers were added to RNA containing vials as specified in Table. Both random hexamer and oligod(T) were used in equal volume in contrast to the manufacturers protocol. The mixture was shortly centrifuged and heated to 65 ºC for 5 min of incubation. The vials were immediately placed on ice bath for swift chilling. Then, the remaining reagents in Table were added to the mixture and gently centrifuged. The initial incubation was done at 25 ºC for 5 min. The main incubation was performed at 42 ºC for 60 min. The activation step was achieved through heating to 70 ºC for 5 min. Reaction products were stored in deep freezer (-20 ºC) for short-term storage.

Table 2.14. The list of the reagents used for first strand cDNA synthesis..

Reaction Components	**Volumes**
Template RNA (DNAse I treated)	11 µL
Random hexamer	0.5 µL
Oligod(T) primer	0.5 µL
5X Reaction Buffer	4 µL
Ribolock RNase inhibitor (20 U/ µL)	1 µL
dNTP mix (10 mM)	2 µL
RevertAid M-MuLV RT (200U/ µL)	1 µL

2.13.2 Amplification of candidate effectors from cDNA templates

The chosen candidate effectors were Pstg_10917, Pstg_11651 and Pstg_13730. The amplification of these genes from cDNA was achieved with designed primers listed in Appendix B. Two product for each gene were expected: i) gene having N-terminus signal peptide region and ii) gene without SP. The cDNAs of 24 hpi, 72 hpi and 10 dpi ASI samples were mixed in equal volume in hope of providing a template with

better success rate for amplification. The PCR amplification was performed with Q5 high fidelity DNA polymerase using the reaction volumes listed in Table 2.6 and the parameters described in Table 2.7. No amplificants were detected for Pstg_13730; the remaining candidates were successfully amplified. The reaction products were stored in deep freezer (-20 ^{o}C).

2.13.3 Cloning to pK7FWG2 vector with Gateway cloning

The pENTR/D-TOPO cloning was performed using the PCR products generated in the previous section. The protocol was same as in Table 2.8 and pENTR/D-TOPO cloning in section 2.11.3. The reaction products were transformed with heat shock method into *E.coli* Top10 competent cells prepared as mentioned before. After recovery phase, the bacteria were plated on LB agar with selective antibiotics, Kanamycin (50 µg/mL). Next day, the positive colonies were checked for presence of the gene of interest by means of colony PCR described in Table 2.9 and 2.10. The verified clones were frozen in liquid nitrogen (20% glycerol) and stored in deep freezer (-80 ^{o}C).

The plasmid isolations of pENTR-Pstg10917-ΔSP (without SP), pENTR-Pstg10917 (with SP), pENTR-Pstg11651-ΔSP (without SP) and pENTR-Pstg11651 (with SP) were conducted using QIAGEN Spin Miniprep Kit (Cat # 27106, Lot # 139315519) as mentioned in Section 2.11.3. The plasmid quality was validated using micro spectrophotometer (NanoDrop). The LR cloning was achieved for each effectors using pK7FWG2 destination vector. The protocol was accomplished as in Table 2.11. The reaction products were transformed into *E.coli* Top10 competent cells. The positive clones were selected using LB agar with Spectinomycin (100 µg/mL). The verifications were achieved by performing colony PCR (previously defined in Table 2.9 and 2.10) and by sequencing plasmids using their corresponding primers (Sentegen Company).

2.13.4 Transformation to *A. tumefaciens*

The verified plasmids (pK7FWG2-Pstg10917-ΔSP, pK7FWG2-Pstg10917, pK7FWG2- Pstg11651-ΔSP and pK7FWG2-Pstg11651) in Section 2.12.3 were isolated and transformed into *A. tumefaciens* GV3101 strains with electroporation.

Preparation of electrocompetent *A. tumefaciens* GV3101 cells

Single colony was grown using streak plate technique on LB agar with antibiotics (Tetracycline, 5 µg/mL and Rifampicin, 10 µg/mL) at 28 ^{o}C. The single colony was transferred into 4 mL of LB medium with Tetracycline (5 µg/mL) and Rifampicin (µg/mL) and incubated at 28 ^{o}C for 2 days at 200 rpm. The grown cells (1 mL) were transferred into bigger sterile flask with LB medium (100 mL) with corresponding antibiotics. The incubation was applied overnight until the optical density of the cells reached to 0.5-0.7 (A_{600nm}). The cells were divided into two sterile falcon tubes (50 mL) and placed in ice bath for 30 min. The cooled cells were centrifuged at 3500 rpm for 15 min. at 4 ^{o}C. The supernatant was removed and the pellet was suspended by pouring 50 mL of ice-cold glycerol (10 %). The centrifugation and washing step were repeated once again. Finally, the cells were resuspended in 200 µL of GYT medium (0125% yeast extract, 10 % glycerol, 0.25 tryptone). The consequential electrocompetent cells were divided into sterile 1.5 mL Eppendorf tubes by adding 50 µL each. The vials were stored in deep freezer (-80 ^{o}C).

Electroporation/Transformation of pK7FWG2 vectors into *A. tumefaciens* GV3101

The plasmid DNA (100-200 ng) of each candidate effectors (pK7FWG2-Pstg10917-ΔSP, pK7FWG2-Pstg10917, pK7FWG2- Pstg11651-ΔSP and pK7FWG2-Pstg11651) was mixed gently with electrocompetent *A. tumefaciens*. The mixtures were transferred into sterile pulsar cuvette and nested on an ice bath for 10 min. The cuvettes

were electrocuted at 25 μF, 2.2 kV range using an electroporator (ThermoScientific, Celljject duo). The transformed cells were nursed with adding 1 mL of LB medium. The incubation was applied at 28 ºC for 1-2 hours at 200 rpm. The selection was done by LB agar with Tetracycline (5 μg/mL), Rifampicin (10 μg/mL) and Spectinomycin (100 μg/mL). After two days of incubation at 28 ºC, positive colonies were observed and colony PCR was applied for validation.

2.14 *A. tumefaciens* mediated cell death assays

Agrobacterium tumefaciens GV3101 carrying pK7FWG2-Pstg10917-ΔSP, pK7FWG2-Pstg10917, pK7FWG2- Pstg11651-ΔSP and pK7FWG2-Pstg11651 were assayed for their ability to suppress cell death induced by various elicitors. pTRBO-GFP (Dr. Sophien Kamoun, The Sainsbury Laboratory, Norwich, United Kingdom) and pK7FWG2-SP-GFP (Dagvadorj et al., 2017) were chosen as negative controls for the experiments. The elicitors for cell death were Inf1 (Kamoun et al., 1999) and PstSCR1 (Dagvadorj et al., 2017). *P. syringae* DC3000 was also used to trigger cell death on tobacco (*N. benthamiana*) as well.

The bacteria (*Agrobacterium tumefaciens* GV3101 carrying pK7FWG2-Pstg10917-ΔSP, pK7FWG2-Pstg10917, pK7FWG2- Pstg11651-ΔSP and pK7FWG2-Pstg11651) were plated on LB agar with selective antibiotics (previously described above) and incubated at 28 ºC for two days. Similarly, *A. tumefaciens* carrying pGR106-INF1 and pJL48-PstSCR were grown on LB agar with Kanamycin (50 μg/mL). *P. syringae* DC3000 was incubated on KB agar with Rifampicin (25 μg/mL) at 29 ºC. The grown cultures were scratched from the surface of their corresponding agar plates with the aid of a sterile pipette tip and suspended in 1 mL of sterile, distilled water. The centrifugation was applied at 4000-5000 rpm for 2 min at room temperature. The supernatant containing the growth medium was discarded. Washing and spinning steps were reiterated again. At final step, *A. tumefaciens* carrying effector candidates and negative controls were resuspended in agroinduction media with OD value of 0.4.

The infiltration was achieved using a sterile, needless syringe via application through the bottom side of the leaf blades of *N. benthamiana* (4-6 weeks old) plants. The infiltration areas (size of small coin) were marked clearly.

The elicitors were challenged after 24 hpi as the reports claimed for better efficiency at that period (Ramachandran et al., 2016) (Cheng et al., 2017). The ODs for the elicitors were 0.4 for Inf1, 0.3 for SCR1 (suspended in agroinduction media) and 0.02 for *P. syringae* DC3000 (DC3000 was suspended in 10 mM $MgCl_2$ solution as in section 2.11.4). The infiltration was performed at the same spot of the ring and without exceeding borders of the previous penetration area. The plants were dried on the bench tops for a while. Subsequently, they were placed in the growth chamber and kept in their regular growing parameters aforementioned in section 2.1. The results for DC3000 were noticeable at 2^{nd} day whereas the cell deaths were apparent at 4^{th} day for both Inf1 and SCR1 challenge. The plants were photographed under both daylight and UV light.

2.15 Subcellular localization experiments

Agrobacterium tumefaciens GV3101 strains carrying pK7FWG2-Pstg10917-ΔSP, pK7FWG2-Pstg10917, pK7FWG2- Pstg11651-ΔSP and pK7FWG2-Pstg11651 constructs were subjected to the microscopy analysis. Briefly, the bacteria were grown on LB agar with Rifampicin (10 µg/mL) and Spectinomycin (100 µg/mL) at 28 ^{o}C for 2 days. The cells were harvested and washed as abovementioned procedure in Section 2.13. The optical density (OD) was accustomed to A_{600nm}: 0.4 using the agroinduction medium. The *N. benthamiana* leaves (4-6 weeks old) were infiltrated with *A. tumefaciens* carrying the constructs with a needless syringe. The assayed areas were noticeably marked with a permanent marker. The plants were dried and put back in the growth chamber programmed to normal growth conditions (see section 2.1). The results were taken at second and third day. Small, thin layers (2-3 mm in length) of leaf pieces were cut from the remote region of the infiltration scar with a sterile scissor.

The bottom side of the leaf blade was placed as the up position on a microscope slide treated with tap water. The observations were done using Leica microscope at 40X magnification using GFP filters. The confocal analysis was conducted at Imperial College, London using a confocal microscope (Leica 385 TCS SP5). The wavelengths chosen during imaging process were excitation at 488 nm and emission at 495–550 nm for GFP fluorescence to visualize the effector localization. The excitation and emission at far infrared (>800 nm) range was used for the visualization of the autofluorescence of the chloroplasts, individually. The merged pictures were recorded for each effectors and chloroplasts to be able to differentiate the plastid localizations with better resolution.

CHAPTER 3

RESULTS AND DISCUSSION

The outcomes of the studies are separated in three parts to provide better comprehension for the readers of the thesis. The pilot studies are presented in the Part I. It includes the experiments conducted on previously, cloned and investigated candidate effectors by Akkaya research group. The main drive to examine these effectors preliminary was to analyze the assays to screen the activity of the effector candidates while the transcriptome sequencing and the micro array studies was progressing. The initial strategy was to investigate the candidate effectors via *P. fluorescens* EtHAn mediated WIA experiments to characterize the effectors in their natural hosts. The results were presented and discussed in Section 3.1 and 3.2.

The results of the sequencing and de novo assembly of the transcriptome of *Pst* along with the list of the DEGs during the compatible and the incompatible interaction were reported in Section 3.3-3.9. This part also includes the core analysis to construct secretome repertoire from PstDEGs and *in silico* characterization practice. The outcome of this analysis serves as a pioneer list of candidate effectors generated from *Pst* Turkish races. Moreover, the micro array experiments were included and discussed in the Section 3.9.

Three candidate effectors predicted to target the chloroplasts or the mitochondria of the host were further analyzed for their effect on the plant cells. The studies to determine biological functions of the effectors are incorporated in Section 3.10 and 3.11.

3.1 *P. fluorescens* EtHAn mediated wheat infiltration assays

The results of *P. fluorescens* EtHAn Mediated Wheat Infiltration Assays (WIA) were listed as a table (Table 3.1). If the marked region of infiltration sides were dominantly showing tissue collapse exposing yellow to brown color, the score was assessed as '1'and if not, the score was '0'. The controls were shown no signs of cell death as expected (Figure 3.1 and 3.2). The scores were assessed for their significance using 'Fischer's exact test'.

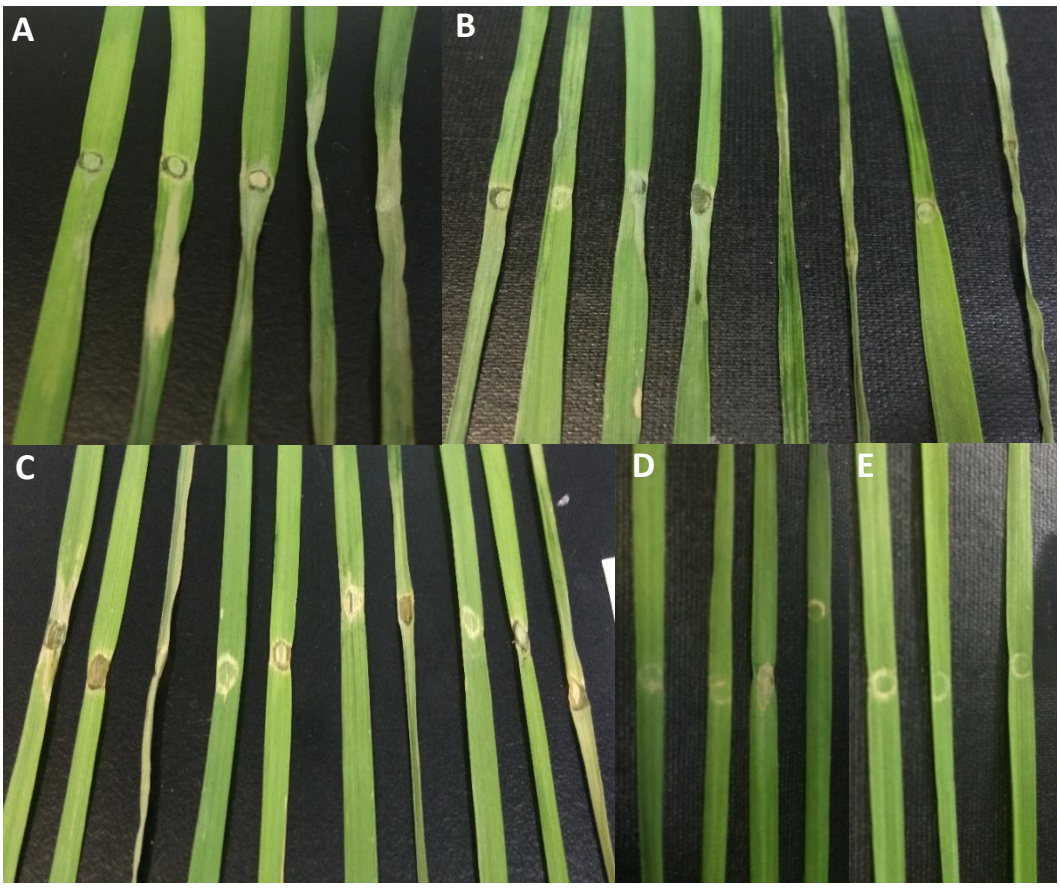


Figure 3.1. Cell death observed on wheat cultivar, Kalyansona (YR2) ($p<0.05$, Fisher Exact Test). *P. fluorescens* EtHan carrying **A.** pEDV6-PstHa12J12, **B.** pEDV6-PstHA15N21, **C.** pEDV6-PstHa12H2, **D.** *P. fluorescens* EtHan (Control) and **E.** MgCl2 (Control). Photos were recorded at 4^{th} dpi.

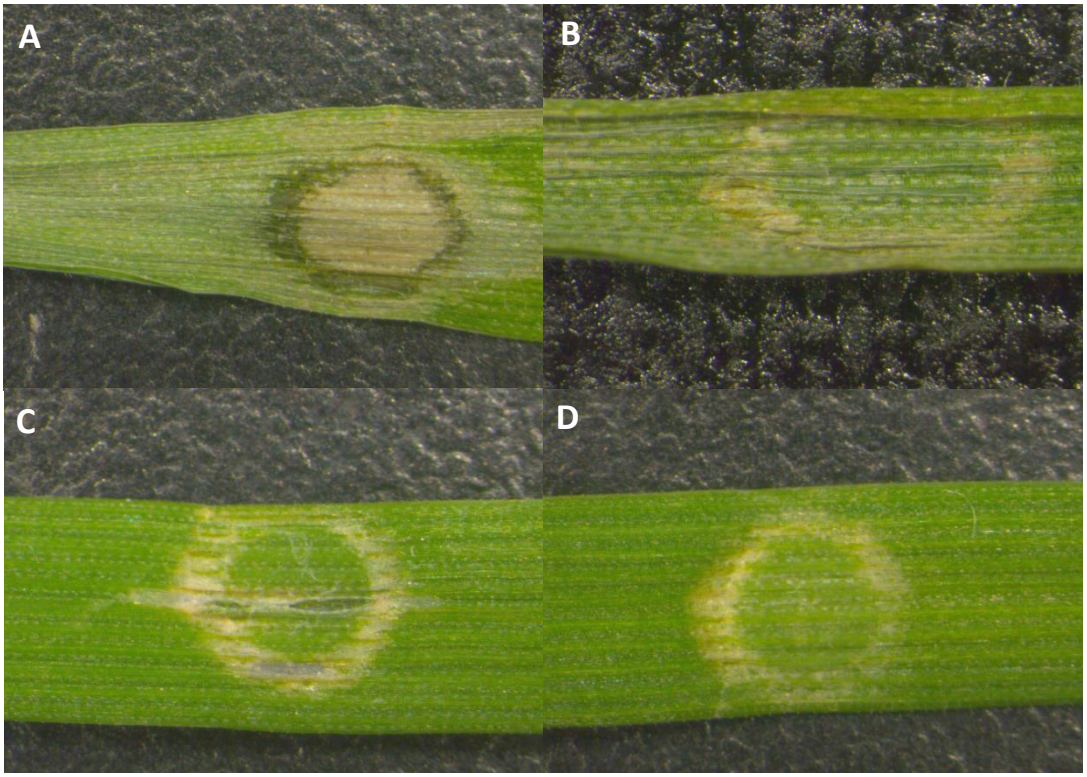


Figure 3.2. Closer look on the cell death region observed on wheat cultivar, Kalyansona (YR2) ($p<0.05$, Fisher Exact Test). *P. fluorescens* EtHan carrying **A.** pEDV6-PstHa12J12, **B.** pEDV6-PstHA15N21, **C.** pEDV6-PstHa12H2, **D.** *P. fluorescens* EtHan (Control) and **E.** $MgCl_2$ (Control). Photos were recorded at 4th dpi.

The outcomes of each effectors were correlated to specific YR resistance genes. However, no certain decision is made due to arbitrary cell death results. For example, all YR2 bearing cultivars respond a cell death for all effectors but PstHa12H12 exhibits irregular number of response (See Table 3.1). Moreover, Pst12J12 shows positive results in two YR18 bearing cultivars whereas it displays no symptoms in Super Kauz. It is possible that lack of uniformity in the results could be attributed to the expression differences, to the scarcity of the number of the effector that could successfully translocated into the hots, to the cellular oscillations due to wounding and to the complex nature of multiple resistance alleles. Hence, we decided to replicate the results using wheat YR differential lines more isogenic to each other and differ only on the type of the resistance genes.

Table 3.1. The cell death counts observed on wheat differential lines after WIA.

YR Differential Lines	12J12	15N21	12H2	EtHAn	$MgCl_2$
Siete Cerros T66 Yr2	4/4*	6/6	3/9	0/3	0/3
Avocet-Yr18	5/8	2/6	0/5	0/3	0/3
Avocet-YrSp	3/5	3/8	4/8	0/3	0/3
Pavon F76 Yr29+Yr30+1	0/5	0/6	0/5	0/3	0/3
Seri M 82 (Yr2, Yr9, Yr29, Yr30, +)	5/5	7/8	1/7	0/3	0/3
Super Kauz (Yr9, Yr27, Yr18)	0/5	0/8	0/10	0/3	0/3
Polmer 2.1.1	0/5	0/8	0/7	0/3	0/3
Avocet-YrCv	3/5	0/7	0/8	0/3	0/3
Avocet-Yr*3/Altar84/Asqu OP(Yr27, Yr18)	5/6	6/8	0/10	0/3	0/3
Avoct-Yr*3//LALBMONO1/PVN	3/4	0/6	1/7	0/3	0/3
Avocet-Yr*3/Pastor (YR?)	4/5	7/10	8/9	0/3	0/3
Jupateco R-Yr18	3/5	2/7	0/7	0/3	0/3
Kalyansona-Yr2	5/5	8/8	5/10	0/4	0/3

* Out of four leaves, all four of them display symptoms of cell death.

The attempts to replicate the findings collected in Table 3.1, was failed several times. The *P. fluorescens* EtHAn produced cell death response while the expectation was to observe no symptoms (See Table 3.2). The other trials were resulted in similar outcomes with cell deaths in negative controls (Data not shown). We hypothesized that the cell deaths instigated from irregularity in viable cell counts during minimal media induction.

Table 3.2. The replication of WIA assay on Avocet near isogenic lines.

YR Type	$MgCl_2$	EtHAn	PstHa12H2	PstHa15N21	PstHa12J12
27	**0/4**	**8/8**	**2/7**	**0/1**	**2/3**
1	**0/2**	**7/7**	**0/6**	**1/3**	**0/5**
5	**0/3**	**3/3**	**0/3**	**0/2**	**0/2**
6	**-**	**6/6**	**0/6**	**0/4**	**0/5**
7	**0/3**	**4/4**	**0/8**	**-**	**0/1**
8	**0/3**	**6/6**	**0/2**	**2/4**	**0/2**
9	**-**	**1/1**	**0/1**	**4/4**	**3/5**
10	**0/1**	**7/7**	**0/2**	**1/4**	**1/6**
11	**0/4**	**2/2**	**0/1**	**4/4**	**1/3**
12	**-**	**3/3**	**1/5**	**0/5**	**0/2**
15	**0/2**	**2/2**	**0/5**	**0/5**	**1/3**
17	**0/5**	**2/2**	**0/4**	**5/5**	**0/4**
18	**0/1**	**8/8**	**0/4**	**1/1**	**0/3**
24	**1/4**	**2/2**	**-**	**1/4**	**0/4**
26	**0/3**	**7/7**	**0/2**	**2/2**	**0/4**
SP	**0/2**	**7/7**	**3/4**	**1/3**	**2/4**
SK	**0/6**	**7/7**	**5/7**	**1/3**	**5/6**
YRA	**0/2**	**4/4**	**0/3**	**0/2**	**0/4**
-(S)	**0/1**	**4/4**	**0/3**	**0/4**	**3/5**
2	**0/1**	**4/4**	**0/4**	**-**	**0/5**

We re-designed the experiments by removing the step of minimal media induction. The bacteria were directly acquired from the KB agar as in the protocol of *A. tumefaciens* infiltration. This approach was previously reported in Liu et al., 2016. The optical density values of 1.0, 1.5 and 2.0 were used for WIA assays. None of the negative controls exhibited a sign of cell death (Table 3.3). In addition, no cell death was observed for the candidate effectors as well (Figure 3.3). We concluded that this

new procedure is efficient to study the function of the candidate effectors on wheat cultivars.

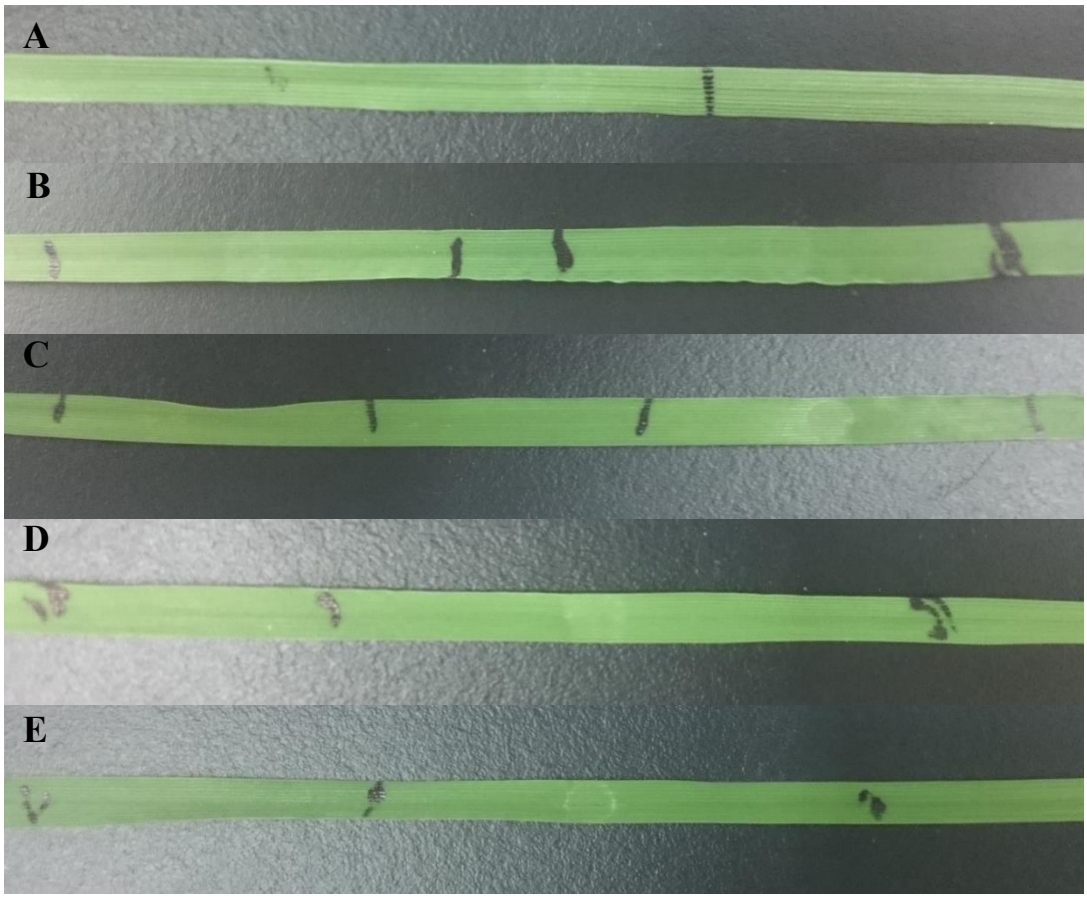


Figure 3.3. The photos of the final WIA assay trial on wheat cultivar, Kalyansona (YR2). **A:** *P. fluorescens* EtHan (control), EtHAn carrying **B:** pEDV6-GFP (Control), **C:** pEDV6-PstHa12J12, **D:** pEDV6-PstHA15N21, **E:** pEDV6-PstHa12H2. Photos were recorded at 4th dpi.

Table 3.3. The results for the final trial of WIA using the effector candidates on wheat cultivars.

YR Type	$MgCl_2$	EtHAn	PstHa12H2	PstHa15N21	PstHa12J12
AvocetS	**0/2**	**0/5**	**0/4**	**0/4**	**0/4**
AvocetYR10	**0/2**	**0/5**	**0/4**	**0/4**	**0/4**
Kalyansona (YR2)	**0/2**	**0/5**	**0/4**	**0/4**	**0/4**

3.2 *P. fluorescens* EtHAn mediated suppression assays

The candidate effectors were tested to scrutinize their ability to suppress PTI dependent defense. To do that, we exploited the potential of *P. fluorescens* to stimulate PTI related defense in the host, *N. benthamiana*. It was reported that *P. syringae* DC3000 related HR or cell death response was halted upon pretreatment of *P. fluorescens* on the same infiltration zone (Chakravarty et al., 2009). Hence, if an effector is able to suppress PTI, DC3000 could successfully induce HR in *P. fluorescens* treated zone. The results for the assay are displayed in Figure 3.4 and 3.5.

Figure 3.4. *P. fluorescens* T3SS delivered PstHa12J12 area is challenged with *P. syringae* DC3000 on *N. benthamiana*. (3 Replicates, at 72 hpi) **P:** *P. fluorescens* EtHAn **G:** pEDV6-GFP, **E1:** pEDV6-PstHa12J12

Figure 3.5. *P. fluorescens* T3SS delivered PstHa15N21 and PstHa12H2 area were challenged with *P. syringae* DC3000 on *N. benthamiana*. (3 Replicates, at 72 hpi) **G:** pEDV6-GFP, **E2:** pEDV6-PstHa15N21 and **E3:** pEDV6-PstHa12H2

None of the candidates were displayed the symptoms of PTI suppression in overlapped area. The expectation was to see cell death due to the suppression in the overlay, which resulted in DC3000 colonization and development of HR. However, the candidate effectors are not involved in PTI suppression.

We investigated the candidates' contribution in the natural host in terms of the suppression of cell death. In similar manner, we performed the same protocol on following wheat cultivars: Avocet S, Avocet YR10 and Kalyansona (YR2). The findings were presented in Figure 3.6. We observed no clear signs of suppression in the natural host.

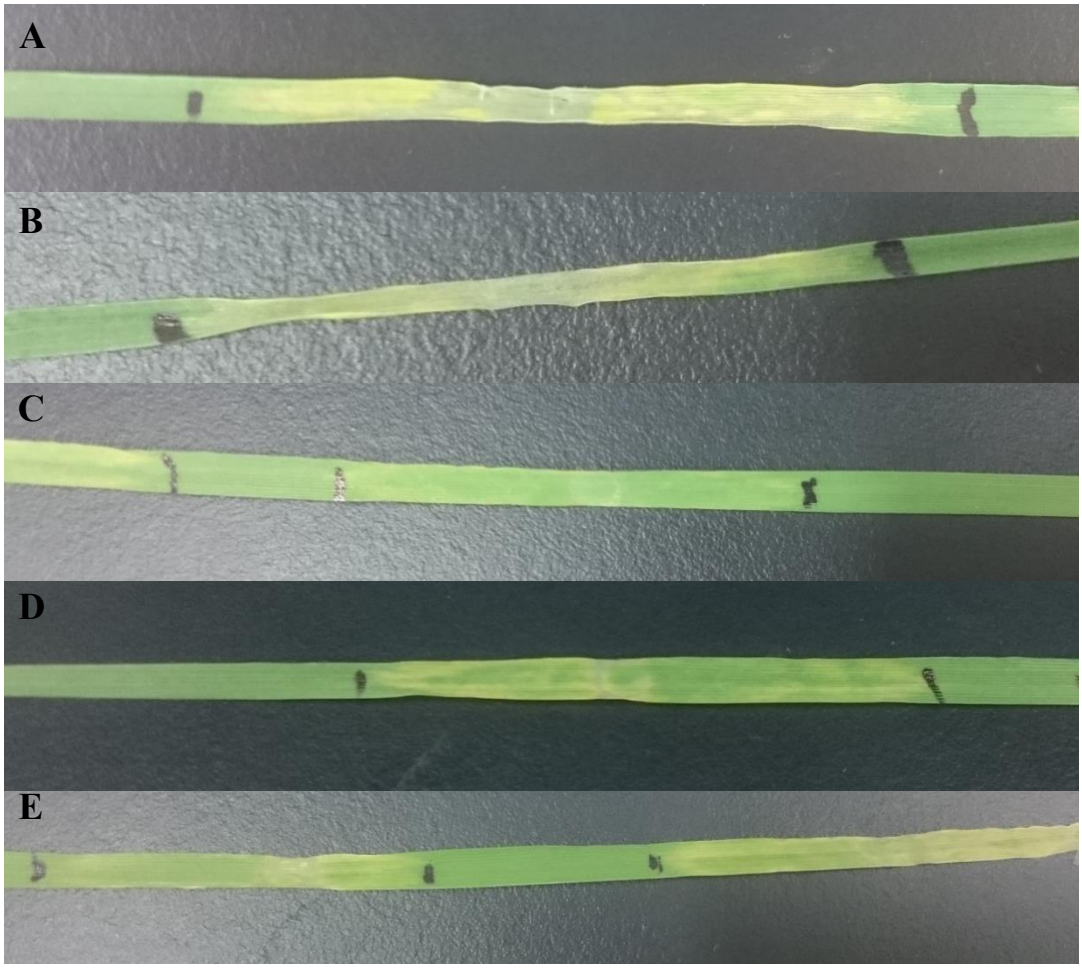


Figure 3.6. *P. fluorescens* T3SS delivered candidates and *P. syringae* DC3000 on wheat cultivar, Avocet S (72 hpi). **A:** *P. syringae* DC3000, **B:** DC3000+GFP (*P. fluorescens* EtHAn pEDV6-GFP as negative control), **C:** DC3000+PstHa12J12, **D:** DC3000+PstHa15N21 and **E:** DC3000+PstHa12H2.

These results suggest that the T3SS mediated effector delivery is useful to study effectors in multiple hosts. However, optimizations are required to screen large number of effectors. The effectors used in pilot studies did not show any reproducible results. Hence, A. tumefaciens mediated transient gene expression on N. benthamiana is chosen to verify the candidate effectors upon the discovery through transcriptome sequencing.

3.3 Transcriptome sequencing

The transcriptome analysis was conducted on four different samples of wheat leaves inoculated either with *Pst* pathogen or with mock treatment. Both Pst genes and host genes were analyzed to uncover compatible and incompatible interaction between the pathogen and the host. However, the genes of *Pst* were covered durig this theses. The generated clean reads of Ps inoculated susceptible and resistant cultivars were listed in Table 3.4. The quality of the reads is assessed with Q20 value, which are 98.16 % and 97.90 % respectively. It should be noted clearly that the obtained reads belong to both Pst and wheat. Therefore, the number of the reads could be evaluated as the impact of the compatible and incompatible interaction on both sides. The numbers are significantly lower for AYI (Avocet-YR10_Pst) samples which could be because of dormant stage of *Pst* and HR related cell death on resistant host. Nevertheless, the high quality reads were promising for comparative transcriptome analysis at 10 dpi of the treatment.

Table 3.4. Statistics of collected clean reads of compatible and incompatible interaction.

Samples	ASI (Avocet-S_PST)	AYI (Avocet-Yr10_PST)
Total Clean Reads	9,204,786	6,899,596
Total Clean Nucleotides (nt)	828,430,740	620,963,640
Q20* percentage	98.16	97.90
GC** percentage	51.44	54.04

*Q20 percentage is proportion of nucleotides with quality value larger than 20.

**GC percentage is proportion of guanine and cytosine nucleotides among total nucleotides.

3.4 De novo assembly

The clean reads with high quality were subjected to assembly analysis to obtain full-length genes. De novo assembly strategy allows us to seize new variants of known genes and novel genes. Hence, we aimed to gather as much of knowledge from *Pst*

TR races. The assembly statistics are listed in 3.5. The graphs for the length distribution of assembled contigs and unigenes are presented in Appendix G.

Table 3.5. *De novo* assembly statistics for *Pst* inoculated/infected wheat transcriptome.

	Samples	**Number**	**Length (nt)**	**Mean Length (nt)**	**N50**	**Distinct Clusters**	**Distinct Singleton**
Contig	AYI (Avocet-Yr10-PST)	105,866	23,189,064	219	249	-	-
	ASI (Avocet-S_PST)	128,680	29,848,685	232	272	-	-
Unigene	AYI (Avocet-Yr10_PST)	55,231	18,753,197	340	371	12,956	42,275
	ASI (Avocet-S_PST)	68,568	24,329,132	355	394	13,564	55,004
	All	61,105	26,978,874	442	490	15,016	46,089

Distinct clusters are similar (more than 70%) unigenes, and these unigenes may originate from same gene or homologous gene; whereas distinct singletons represent the unigene come from a single gene.

3.5 Annotation and classification

The assembly results were annotated using several public databases listed below (Table 3.6). Among all assembled unigenes (61105), 26843 of them were annotated against at least one of the public databases. The assembly results were containing both *Pst* and wheat genes; hence, detailed analysis like biological functions etc. were postponed until we separated two organisms. On the other hand, the results against fungus NR classification are presented as Figure 3.7 to provide visual assessment of assembled unigenes.

Table 3.6. Annotation statistics of the assembled unigenes of transcriptome sequencing project.

Sequence file	**NR**	**NT**	**Swiss-Prot**	**KEGG**	**COG**	**GO**	**All**
All-Unigene.fa	24,933	9,945	16,318	16,275	13,278	10,015	26,843

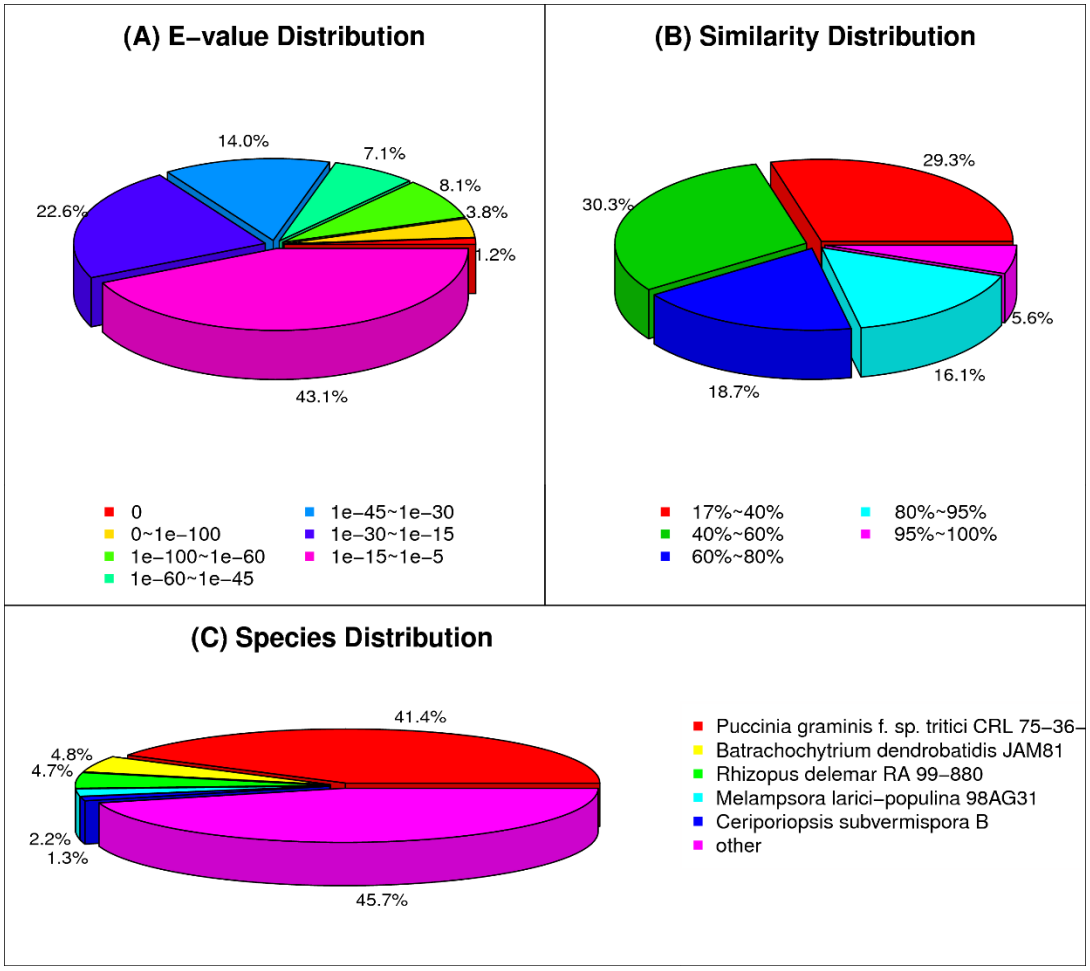


Figure 3.7. Statistics of NR classification of unigenes. **(A)** The E-value distribution of the alignment results of NR annotation. **(B)** The similarity distribution and **(C)** the species distribution of the result of NR annotation.

3.6 Identification and mapping of DEGs

Differentially expressed unigenes (DEGs) were identified using calculations mentioned in Section 2.5. The analysis resulted in 10550 unigenes, which are significantly higher than incompatible group. The distribution of DEGs are given as a graph in Figure 3.8, in which 6851 unigenes are up regulated and 3699 down regulated in compatible interaction compared to incompatible ones.

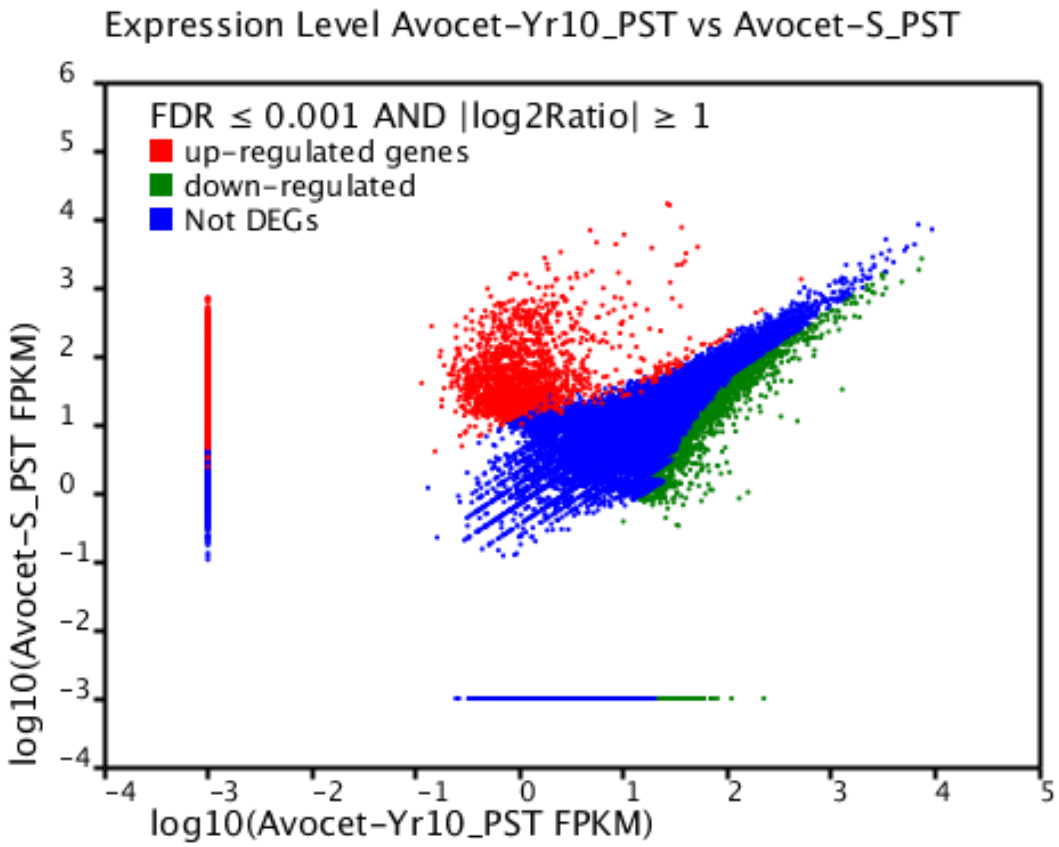


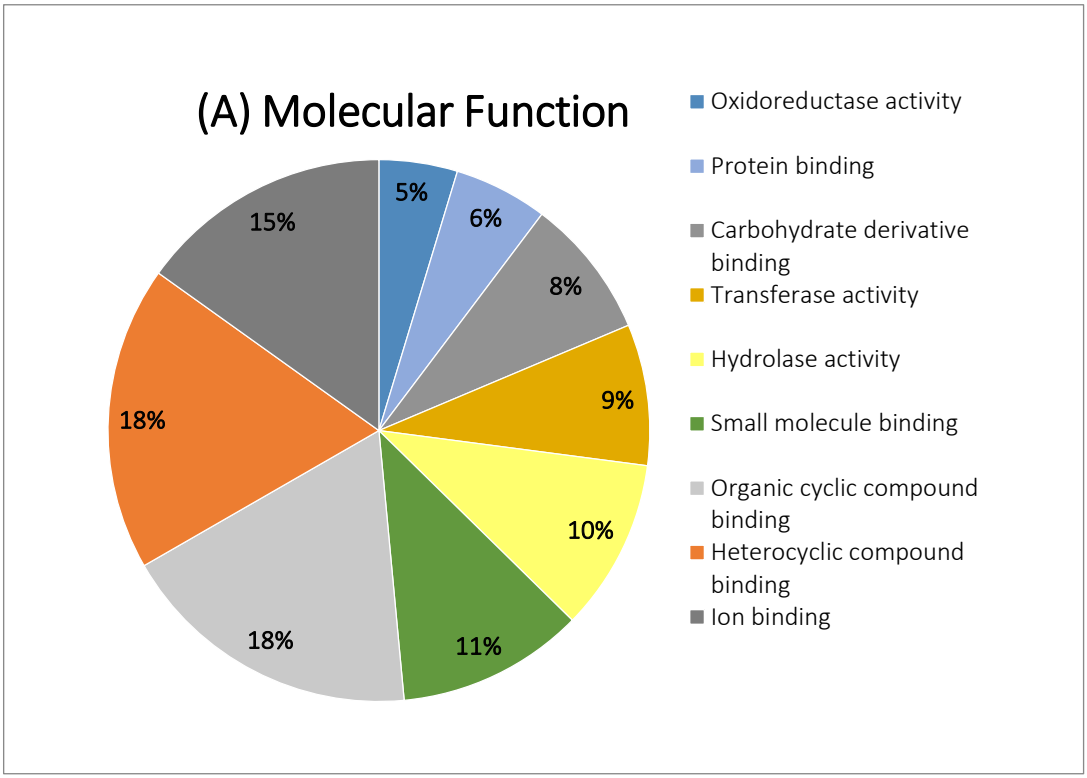
Figure 3.8. Distribution of differentially expressed unigenes (DEGs); 6851 up, 3699 down regulated.

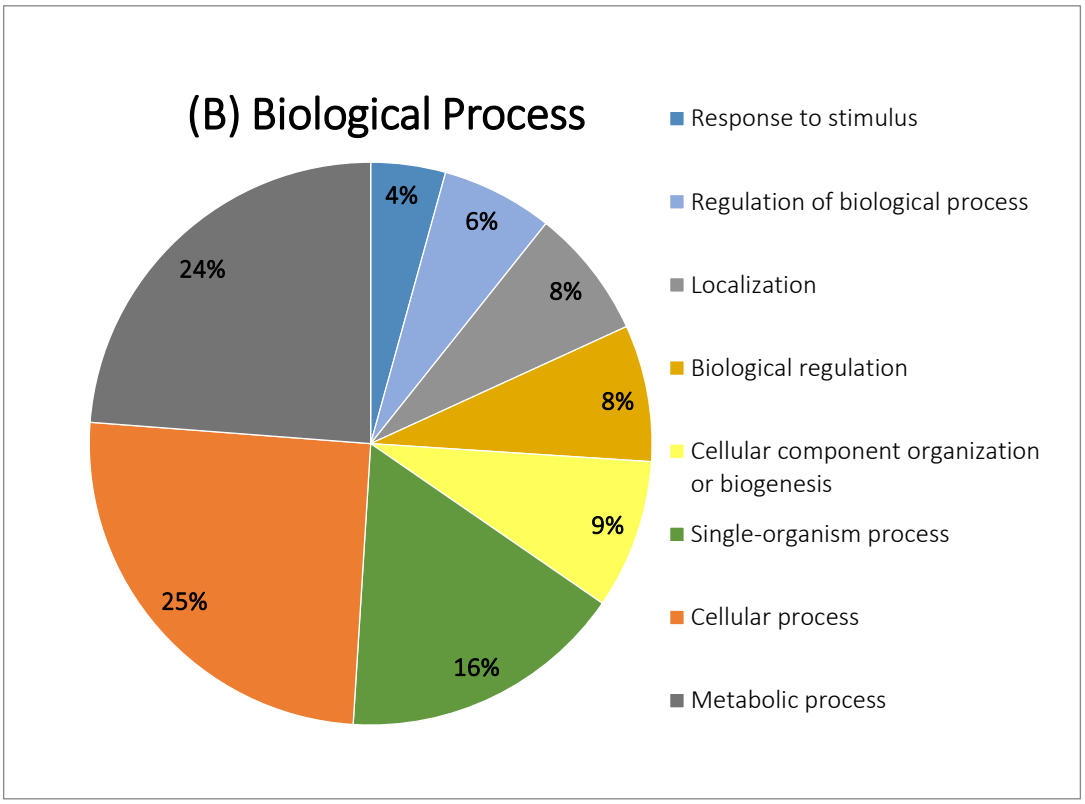
The aforementioned DEGs were belonged to both *Pst* and wheat. We filtered the *Pst* related DEGs (PstDEGs) using a customize strategy described in Section 2.6. To avoid complication, the unigenes with high similarity to both *Pst* and wheat were sorted out based on best hit score. The unigenes that gave the best hit to *Pst* than wheat were categorized as PstDEGs. The major percentage of the PstDEGs was up regulated compared to incompatible interaction. This is an expected finding since biotrophic nature of the pathogen suggests that the genes should be activated at 10 dpi of disease stage. Contrariwise, it should project dormant nature on the incompatible host. We also listed the unigenes that showed no homology to either *Pst* or wheat, were termed

as novel genes (Data not shown). The PstDEGs are evaluated as the main findings of the thesis study and subjected to further studies.

The PstDEGs were functionally annotated using Blast2GO software. The results are presented as graphs of three category: i) molecular function, ii) biological process and iii) enzymatic activity classification (Figure 3.9). The significant amount of PstDEGs was annotated to have catalytic activity (2337). The metabolic activity is expected to reach its peak levels due to biological interface (haustoria) is established and association proceeds continuously at 10 dpi of disease progress in contrast to the dormant and pernicious environment of incompatible interaction. In similar manner, the up-regulated PstDEGs are fitting to variety of biological process shown in the graph (Figure 3.9B). We get curious about the proportion of enzyme classes that are active during compatible interphase. The findings suggest that the hydrolases are the main class of enzymes in PstDEGs (Figure 3.9C). Suppressing the defense, establishing a continuous interface for feeding purpose and manipulating the host seem to require hydrolases at most, while the remaining enzymes are not disposable as well.

The *Pst* pathogen has more than 20000 predicted genes, which is significantly higher than *Pgt* or *Ptt*. Our findings emphasize that at least 6220 of them active at 10 dpi. Moreover, the PstDEGs consist of dominantly genes related to metabolic and biological processes rather than structural ones.


(A) Molecular Function
5%
6%
8%
9%
10%
11%
18%
18%
15%
Oxidoreductase activity
Protein binding
Carbohydrate derivative binding
Transferase activity
Hydrolase activity
Small molecule binding
Organic cyclic compound binding
Heterocyclic compound binding
Ion binding


(B) Biological Process
4%
6%
8%
8%
9%
16%
25%
24%
Response to stimulus
Regulation of biological process
Localization
Biological regulation
Cellular component organization or biogenesis
Single-organism process
Cellular process
Metabolic process

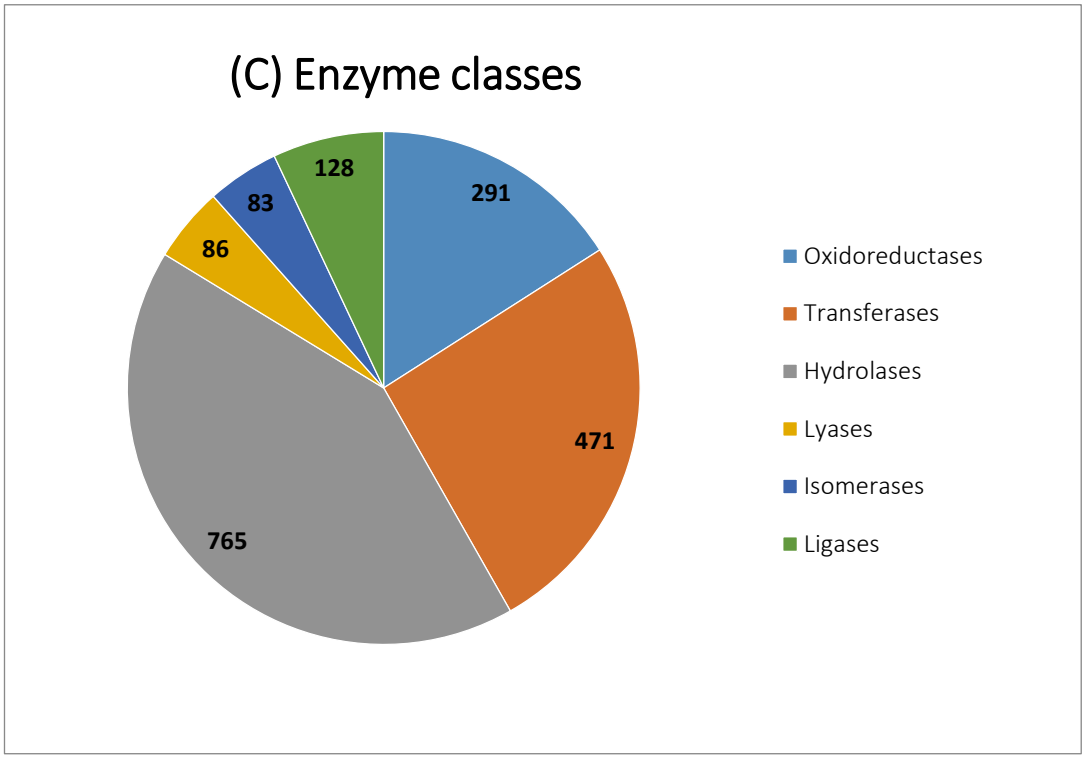


Figure 3.9. Functional annotations of PstDEGs via Blast2Go program. **(A)** Molecular function (Level 3), **(B)** Biological process classification (Level 2) and **(C)** Enzyme class distribution of PstDEGs.

3.7 Differentially expressed small secreted proteins of *Pst* (PstDESSPs)

PstDEGs that generated through transcriptome sequencing is a valuable data elucidates the disease related pathogen genes involved at 10 dpi time point. However, we focus on narrower perspective of *Pst* effectors stimulated at 10 dpi time point of the disease development. To manage that, we constructed a secretome repertoire using PstDEGs through prediction programs. We filtered the small proteins, which are shorter than 300 amino acids in length in their mature form (without signal peptide). Among PstDEGs (6220), we identified 230 of them to be small-secreted proteins by applying definite decision rule described in Section 2.7 (Appendix C-F). These differentially expressed small-secreted proteins are coined up as PstDESSPs.

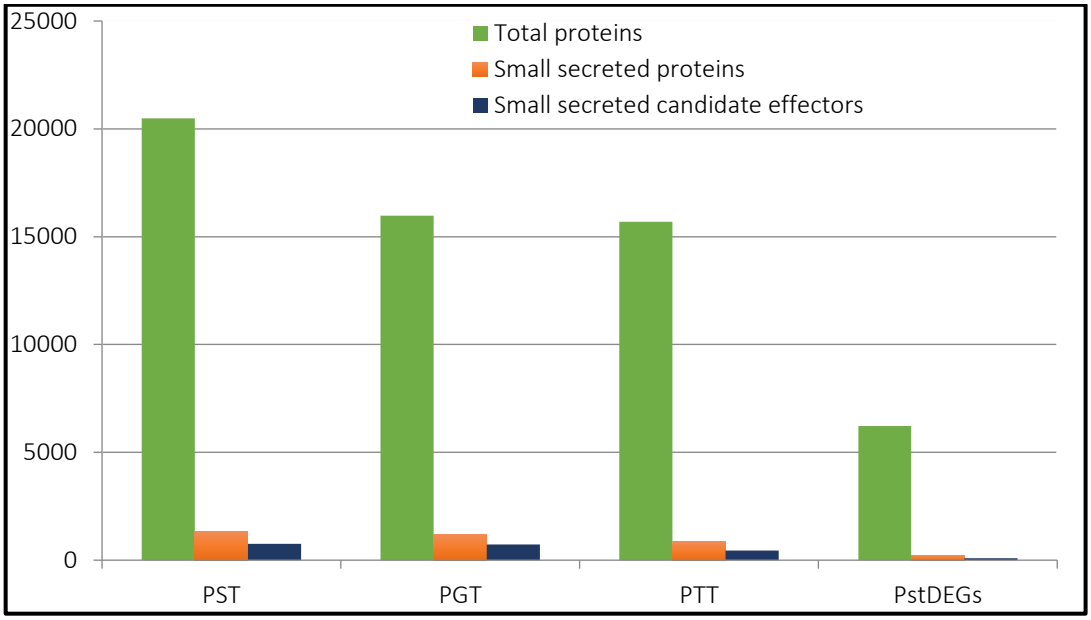


Figure 3.10. The graphical comparison for the predicted small-secreted proteins (SSPs) and small-secreted candidate effectors (SSCEs) of PstDEGs and other Pucciniale proteome.

The same pipeline was applied to all predicted proteome data of *Pst*, *Pgt* and *Ptt* in an attempt to assess the correlation between the regulated genes and the rest. The numbers are presented as a graph in Figure 3.10. *Pst* (20482) has largest proteome repertoire among its relatives *Pgt* (15979) and *Ptt* (15685). However, we forecast 1332 SSPs (6.5 %) for *Pst*, whereas 1211 (7.6 %) and 872 (5.6 %) are the numbers of SSPs predicted for *Pgt* and *Ptt*, respectively. Secreted protein potential of rust genomes were studied extensively in number of studies (Kim et al., 2016) (Xia et al., 2017). We observe the number of SSPs was highest for *Pgt* based on percentages followed by *Pst* and *Ptt*. Here, we found 230 PstDESSPs corresponding to 3.7 % of total PstDEGs. Hence, 3.7 % of PstDEGs are SSPs and involved at 10 dpi of the disease progress. If we correlate the number of PstDESSPs (230) to total number of SSPs (1332) predicted from *Pst* proteome, 17.3 % of the SSPs are regulated in compatible interaction compared to incompatible interaction. As expected, the fraction is quite high. An obligate biotrophic pathogen requires large number of SSPs to establish an interface

and maintain communication with its host because SSPs could translocate into host and apoplastic environments easily. The results are in parallel with the expectations.

The EffectorP program is a machine-learning algorithm developed specially for predicting effector candidates from secretomes (Sperschneider et al., 2017a). We projected small-secreted candidate effectors (SSCEs) from PstDESSPs using EffectorP (v 2.0). Moreover, we performed same analysis on SSPs generated on *Pst*, *Pgt* and *Ptt* (Figure 3.10). The analysis shows 94 (40.9 %) of PstDESSPs are SSCEs. Therefore, high fraction of predicted SSPs is indeed projected to have effector functions involved in disease formation, progress and maintenance. The numbers of SSCEs for *Pst*, *Pgt* and *Ptt* are 755 (56.7 %), 730 (60.3 %) and 447 (51.3 %) in subsequent order. Hence, the real outcome is uniform with projected SSCEs numbers of total SSPs. Hereafter; we can proceed with further characterization of PstDESSPs to deliver the most favorable candidates for the biological function tests.

3.8 *In silico* characterization of PstDESSPs

We characterized PstDESSPs by searching against various databases, past reports, prediction programs etc. and merged the results as a single integrated table to offer overall look on all investigations (Tables in Appendix D, E, F and G). The findings during functional annotations (Figure 3.9c) direct us to reveal enzymatic properties (especially hydrolases) of PstDESSPs. We juxtaposed PstDESSPs through similarity search against public enzyme databases of proteases (MEROPS), lipases (LED), CAZymes (dbCAN) and oxidoreductases (fPoxDB). The findings are displayed in Appendix C and the numbers are given as a graph in Figure 3.11.

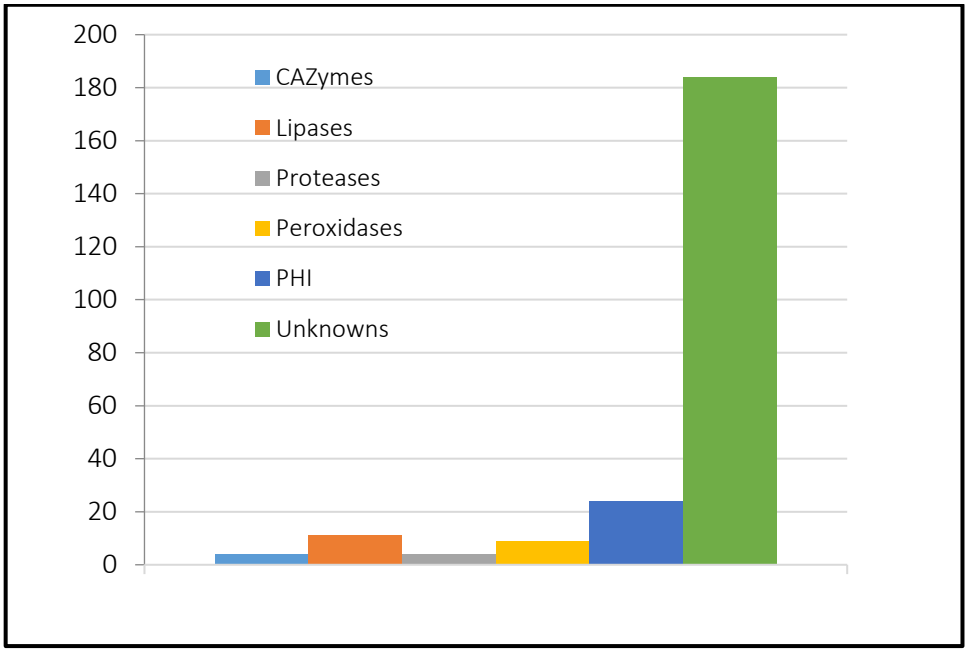


Figure 3.11. Annotated PstDESSPs through similarity search against databases.

The pathogen host interaction database (PHI-base) serves as a catalogue of known proteins to have virulence or pathogenicity attribution curated from published reports (Urban et al., 2015). The PstDESSPs shows high similarity to PHI-base is listed in Table (Appendix E) and Figure 3.11. The remaining PstDESSPs are indicated as unknowns. The majority of the PstDESSPs are unknown for their functions related to pathogenicity as expected from the effector candidates. Previously, it was discussed that an effector candidate should be either unknown or annotated of functions related to pathogenicity (Saunders et al., 2012). The results of PstDESSPs are in parallel with the publications for being the reliable candidates for pathogenicity. Furthermore, we investigated the conserved domains in PstDESSPs with NCBI conserved domain database (CDD). The analysis shows some of the PstDESSPs to bear previously unannotated peptidases, oxidases etc. related domains, lytic transglycolase (DPBB_1) or fungus specific cysteine rich domains (CFEM) (Appendix C).

Subcellular localization predictions programs aid to speculate about an effector candidates for the function or attribution inside host. We use various prediction programs to predict the location inside the host. The mature protein sequence was used because it was assumed that upon secretion, an effector loses its SP and translocate inside host cell via unknown mechanism. The results are displayed in Table in Appendix D. The mitochondria and chloroplast targeting effectors are quite high in proportion. The chloroplast and mitochondria are the main organelles responsible for release of reactive oxygen species (ROS) which could trigger main host defenses including callose deposition, lignin fortification, signal transduction and even cell death itself. Hence, it is tempted to speculate pathogen could mimic host transit peptide sequences to target its organelles in an attempt to interfere host defense. The effector candidates that are destined to apoplastic fluid were forecasted using machine-learning algorithm of ApoplastP program. The effectors in apoplastic fluids have tendency to possess more cysteine residues to ensure their stability. Therefore, we included the cysteine count number of mature PstDESSPs in analysis and marked proteins that bear more than 3 % of cysteine in length.

The previously reported conserved fungal motifs in effectors are examined and included in Table in Appendix C.

The phylogenetic analysis was conducted on PstDESSPs to figure out the similarity levels if exist any. Multiple sequence alignment indicated that three main groups are present among PstDESSPs (Figure 3.12). The conserved motif search was conducted on the branches to identify any conserved or repeated motifs in PstDESSPs. We detected no significant motifs to report. The fungal effectors of plant pathogens are known to lack of conserved motifs (Selin et al., 2016) except the aforementioned motifs. Hence, the findings are consistent with the literature.

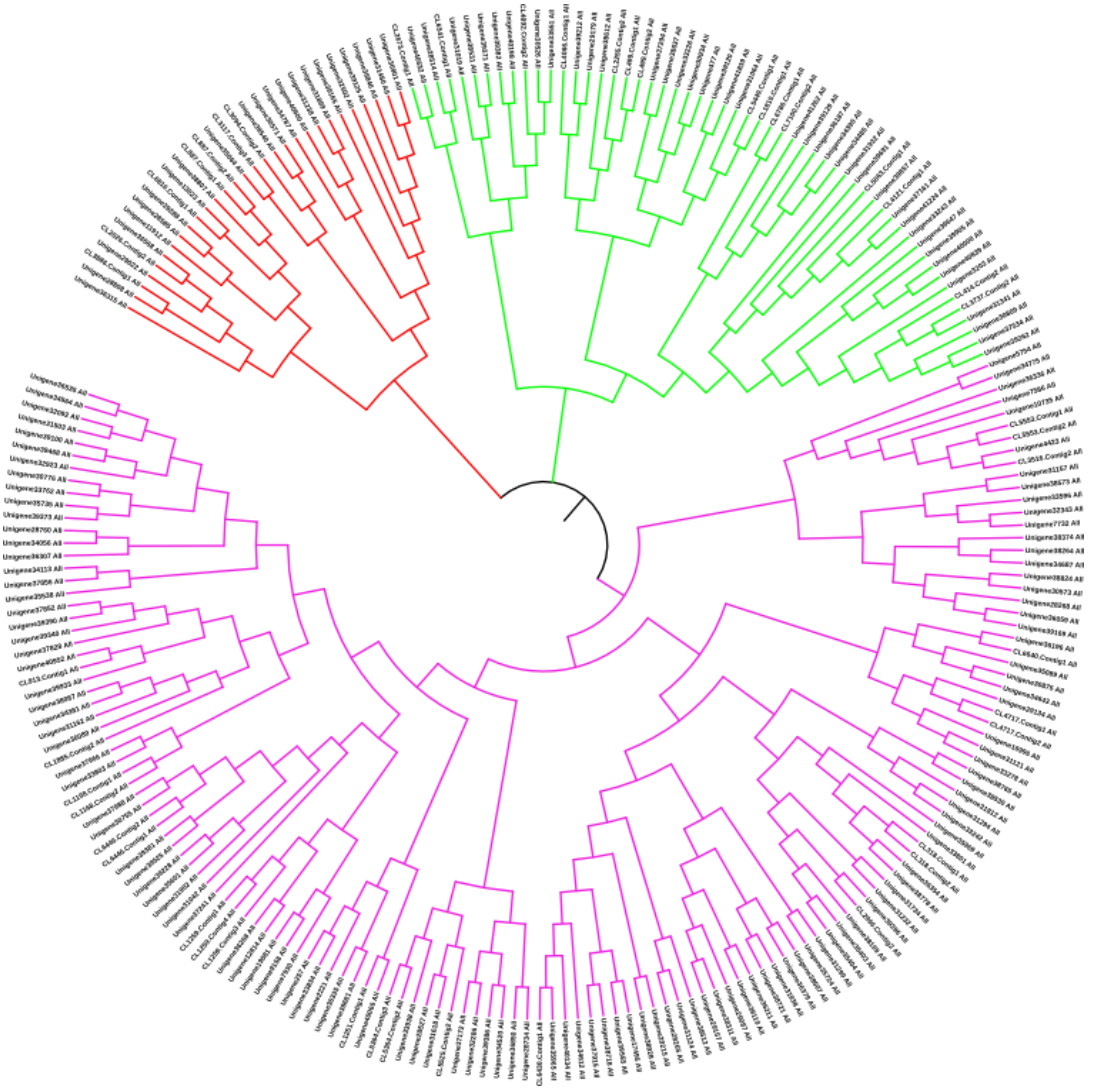


Figure 3.12. The phylogenetic tree constructed among PstDESSPs using sequence similarities. Three main branches were observed at first glance in terms of relationship and marked with different colors.

PstDESSPs are unique to our study of compatible and incompatible interaction at 10 dpi. However, there are several reports belong to different time points of the disease development. Therefore, we compare the PstDESSPs with the reported candidates of 'haustoria specific effector proteins' (HSEPs) and 'infection specific effector proteins' (ISEPs) by Cantu and his colleagues (Cantu et al., 2013). The local blast analysis shows that significant number of PstDESSPs is similar with the HSEPs and ISEPs.

Moreover, some of the PstDESSPs are unique to the 10 dpi time point and are first to be revealed in this study. Additionally, Xia et al. reports a correlation analysis on 7 existing and 7 new sequenced genomes of Pst to uncover Avr candidates for YR resistance genes (Xia et al., 2017). Xia and co-authors compared their candidates with HSEPs and ISEPs of Cantu et al. 2013. The Avr candidates are discarded if they were not matched. However, we exercised the Avr candidates with PstDESSPs and detected novel matches, which was previously neglected on the work of Xia et al. 2017. Clearly, the literature comparisons emphasize the novelty and reliability of PstDESSPs to serve as candidates for biological function investigations.

3.9 Findings of microarray analysis

The microarray analysis conducted on *Pst* TR0997 (one of the race in Pst TR Mix exhibiting same avirulence on YR10) using probes designed based on Pst-78 genome. Early version of Pst-78 genome was available in 2014 whereas final version was published in 2017 (Cuomo et al., 2017). Hence, the probes were designed using early release. The samples collected at 24 hpi and 10 dpi of both compatible and incompatible interaction. The top 10 up-regulated *Pst* genes are displayed as a graph for 24 hpi and 10 dpi separately whereas the remaining are not exhibited in the thesis work.

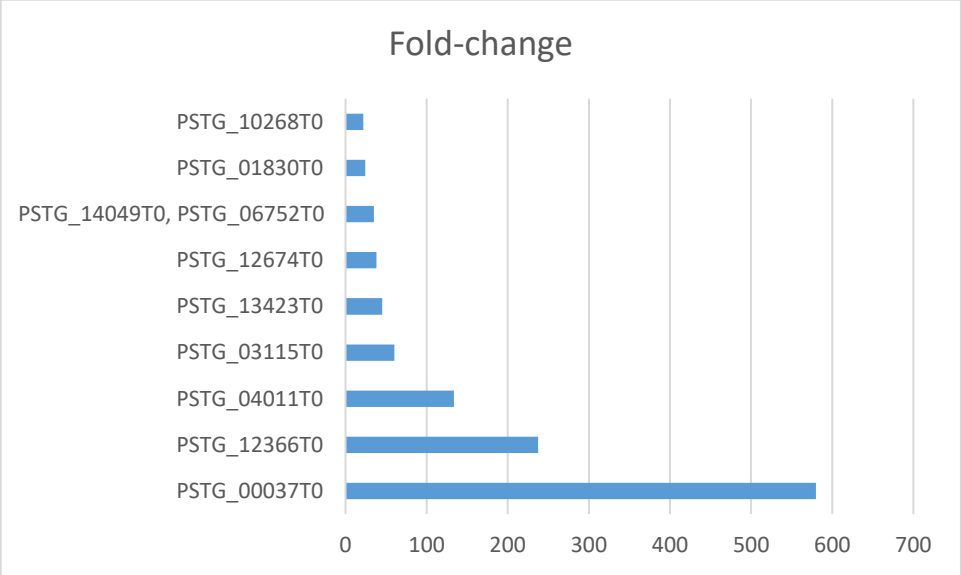


Figure 3.13. The 10 highest upregulated *Pst* genes generated at 24 hpi of the disease progress.

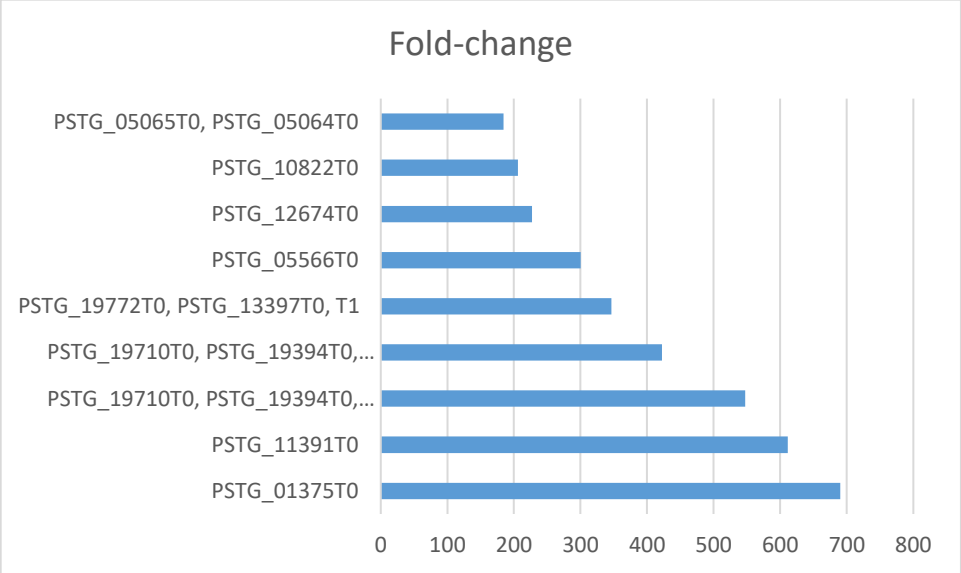


Figure 3.14. The 10 highest upregulated *Pst* genes generated at 10 dpi of the disease progress.

The microarray analysis was not in parallel with the findings obtained in transcriptome sequencing especially for 'absent-present' genes. The genes that are not present in one condition while abundant in other one were not observed in the data. Filtering parameters used to evaluate the microarray results provide statistical significance for detecting differential expression of the genes present in both conditions. Hence, the generated sets of differentially expressed genes mainly belong to metabolic pathways and secreted proteins are absent in the data.

3.10 Cell death suppression assays of candidate effectors

A phytopathogen needs effector proteins in order to suppress defense system while establishing continuous feeding apparatus. Hence, we tested our candidate effectors (Pstg10917 and Pstg11651) whether they can suppress cell death induced by various inducers. INF1 is one of the elicitor of *P. infestans* capable of stimulating programmed cell death (PCD) following expression on *N. benthamiana* leaves (Kamoun et al., 1998). The effector candidates of various phytopathogens were studied extensively for their attribution upon INF1 challenge (Cheng et al., 2017) (Zhao et al., 2018). In similar approach, Pstg10917 and Pstg11651 were investigated for their suppression ability with or without SP in N-terminus. Pstg10917ΔSP-GFP exhibits a suppression pattern against INF1 induced PCD contrasting to the result of Pstg10917-GFP (Figure 3.15). Pstg10917 (Unigene17495) has homologs in *Pgt*, *Ptt* and other Pucciniale species. A distant homolog of Pstg10917 (E-value, $1e^{-14}$) in *Phakopsora pachyrhizi* is PpEC82 that reported as candidate effectors (Link et al., 2014). A recent study demonstrated that PpEC82 has also suppressing ability against BAX induced PCD (Qi et al., 2018).

Pstg11651-GFP displayed a suppression whereas Pstg11651ΔSP-GFP did not halter the response. Fifteen replicates were assayed to increase similar results. Pstg10917ΔSP-GFP produced three complete or dominant suppressions (strong), three moderate suppressions, three weak suppressions and remaining six replicates did

not exhibit any symptom. The Fisher exact test statistic value is 0.0007 for the findings and the result is significant at the statement of significance ($p < 0.05$). Thereby, Pstg10917ΔSP-GFP (Unigene17495) is capable of hampering PCD induced by INF1 elicitor significantly. On the other hand, Pstg11651-GFP only showed one strong and one weak suppression out of fifteen replicates. All negative controls worked smoothly as expected. Hence, Pstg11651-GFP needs more investigation but it did not suppress PCD induced by INF1 in significant manner.

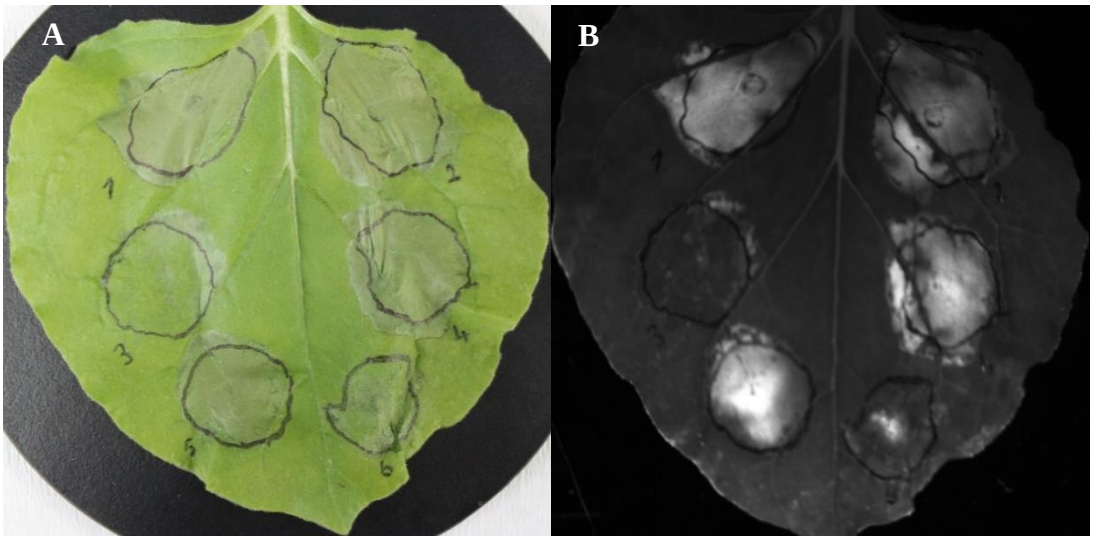


Figure 3.15. Inf1 triggered 'cell death suppression' assay. **(A)** Day light and **(B)** UV light exposure. Candidate effectors are expressed with *Agrobacterium tumefaciens* and challenge with Inf1 at 24 hpi. Cell death area photographed at 4 days after Inf1 expression **(1)** GFP (control), **(2)** SP-GFP (control), **(3)** Pstg10917ΔSP-GFP and **(4)** Pstg10917-GFP, **(5)** Pstg11651ΔSP-GFP and **(6)** Pstg11651-GFP.

The candidate effectors were also tested against *P. syringae* DC3000 and PstSCR1 to understand the suppression occurs against all cell death inducers or else. The effector constructs with or without SP was expressed on *N. benthamiana* leaves and challenged with inducers at 24 hpi on the same infiltration ring. No visible suppression was observed against any elicitor (Figure 3.16). These results suggest that Pstg10917 is effective against only INF1 triggered cell death. It is possible that Pstg10917 only

involves in INF1 mediated pathway of PCD not in general signaling leads to PCD. *P. syringae* DC3000 suppresses PTI response and deliver its effectors inside host cell but HR was triggered due to recognition of effectors. Hence, *P. syringae* DC3000 uses HR dependent PCD. On the other hand, PstSCR1 triggers PCD at the cell surface level; possibly benefitting BAK1 mediated PTI system (Dagvadorj et al., 2017). Pstg10917 did not suppress any of them in significant manner. In literature, the effector candidates are able to suppress all forms of PCD in general but some of them are successful suppressors only for certain cell death response (Ramachandran 2016) (Cheng et al., 2017) (Zhao et al., 2018) (Qi et al. 2018). The finding suggests Pstg10917 involve in INF1 triggered cell death but not *P. syringae* DC3000 mediated ETI and PstSCR1 mediated cell death. Hence, it is possible that Pstg10917 is effective on specific signals or targets of certain defense pathways.

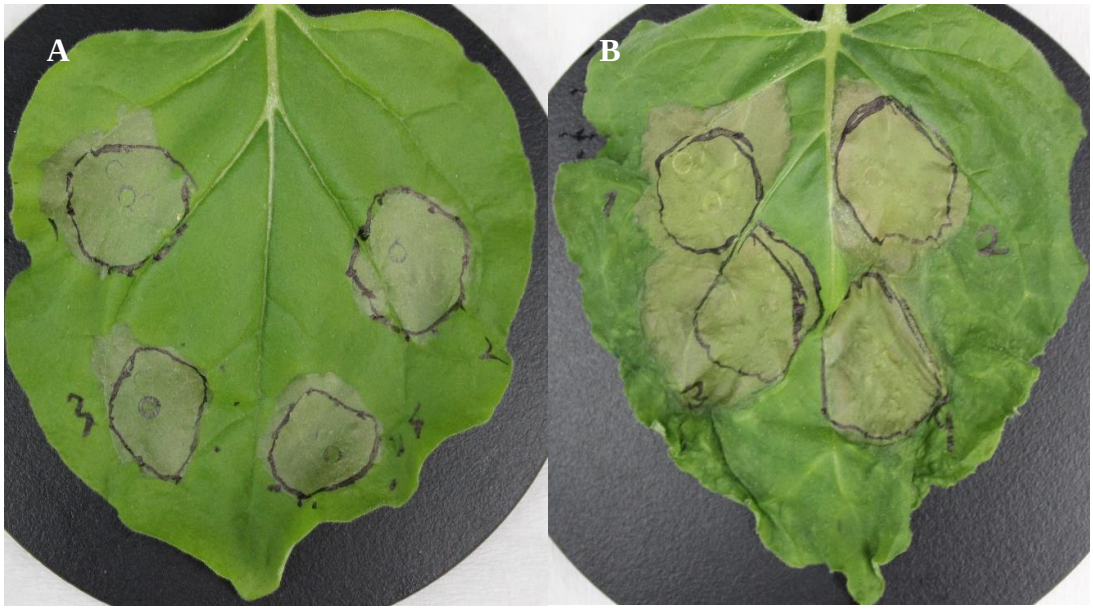


Figure 3.16. *P. syringae* DC3000 (A) and PstSCR1 (B) triggered 'cell death suppression' assay. Candidate effectors are expressed with *Agrobacterium tumefaciens* and challenge with *P. syringae* DC3000 and PstSCR1 at 24 hpi. Cell death area photographed at 4 days after Inf1 expression (1) GFP (control), (2) SP-GFP (control), (3) Pstg10917ΔSP-GFP and (4) Pstg10917-GFP.

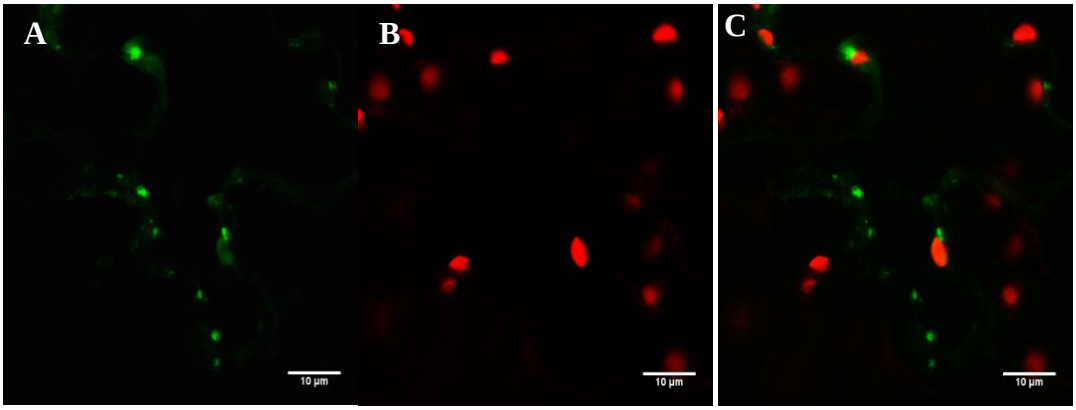
Pstg11651 did not display suppression of *P. syringae* DC3000 and PstSCR1 (Data is not shown). Pstg11651 bear a DPBB_1 domain based on the conserved domain prediction. The DPBB_1 domain of PNPi effector was previously shown to interact with NPR1 protein in order to jam transcriptional regulation (Wang et al., 2016). It was expected for Pstg11651 to suppress cell death but our findings are not suggesting any interference. Pstg11651 may interfere with other defense related response such as callose deposition, hormone signaling or other defense related signaling as PNPi. Hence, Pstg11651 was not characterized in this study but further studies may elucidate its biological relevance.

3.11 Subcellular localization analysis of candidate effectors

The candidate effectors were assayed to investigate their site of localization inside the plant cell. *N. benthamiana* leaves were used for stable expression of the effectors. The effectors with or without SP were fused with GFP tag on their C-terminus. The results were displayed as both fluorescence and confocal microscope images (Figure 3.17 and Figure 3.18). The findings are in union with *in silico* prediction programs. Both of the candidate effectors were projected to bear a transit peptide (TP) sequence following signal peptide (SP) region. Hence, the effectors were predicted to localize in the chloroplasts or mitochondria. Prior to the experiment, it was expected for an effector with SP to localize in apoplastic fluid and effector without SP should target the organelle. However, Pstg10917 seems to target the chloroplasts with or without SP (Figure 3.17). It should be noted that Pstg10917 without SP seems to localize near the chloroplasts but not in the chloroplasts like Pstg10917 with SP. It is clear that such polarized localization on the chloroplast membrane is not likely because of the homogenized patterns of receptors in the chloroplasts. We concluded that the further microscope analysis might be needed to decide the exact localization. Localizer program detected a TP region inside the SP region whereas other programs failed. Pstg10917 with SP might have a TP region inside the SP sequence, which could

elucidate how the effector targets the chloroplasts although it has a secretion signal. It is possible that the hidden TP region leads to chloroplast targeting because of dominancy. Other possible explanation, the effector was secreted to the apoplast but internalize back with an unknown tactic of pathogen independent pathway. Pstg10917 did not possess any of the known motif that validated in effector translocation. Hence, the situation needs to be examined further. It should be also noted that only Pstg10917ΔSP-GFP able to suppress INF1 mediated cell death whereas both constructs reveal same subcellular localization pattern. The overlapping regions are not identical in confocal analysis suggesting an analysis of microscope with higher resolution could enlighten the correct subcellular or even sub-organelle localization.

Pstg10917ΔSP-GFP



Pstg10917-GFP

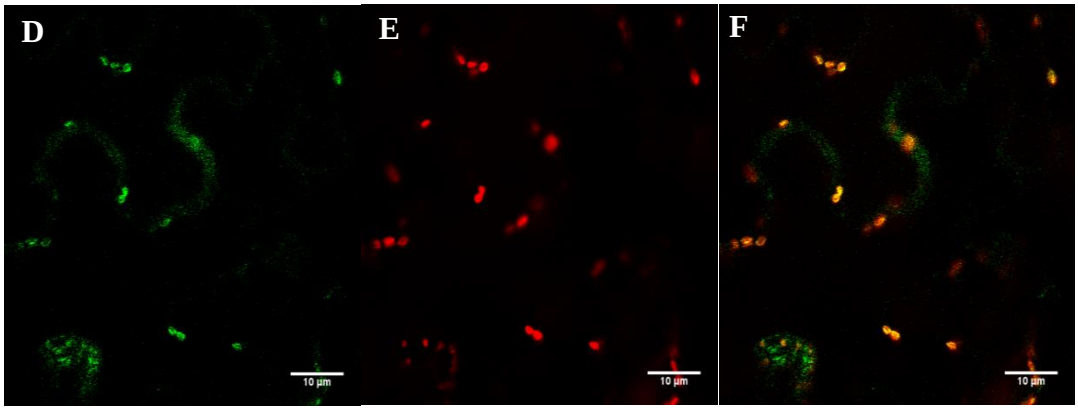


Figure 3.17. Subcellular localization of Pstg10917-GFP (Unigene17495) inside *N. benthamiana* leaves. **(A)** Expression of Pstg10917ΔSP-GFP **(D)** Expression of Pstg10917 **(B)** and **(E)** Autofluorescence of the chloroplasts, **(C)** and **(F)** Overlaid images. Expression on *N. benthamiana* was achieved using *A. tumefaciens*. The leaves were photographed under confocal microscope (Leica 385 TCS SP5) at 2nd day of the expression. The excitation was at 488 nm and the emission at 495–550 nm (GFP fluorescence) whereas the excitation and emission at far infrared (>800 nm) wavelength to visualize chloroplast autofluorescence.

Pstg11651 was predicted to localize in mitochondria. However, the microscope analysis reveals that it targets the host chloroplasts. TP region is required for both mitochondria and chloroplast targeting and prediction programs may fail to identify correct organelle. Studies demonstrated that an effector of poplar rust (*Melampsora*

larici-populina) successfully targets both mitochondria and chloroplast (Petre et al., 2015 and 2016). In parallel, Pst11651 targets the chloroplast not the mitochondria with TP region. Pst11651 shows apoplastic localization pattern meaning it could be successfully secreted. It needs the pathogen itself to be delivered inside the host cell. It is unknown if the effector is cytoplasmic or apoplastic but if it manages to enter the host cell it will target the chloroplasts.

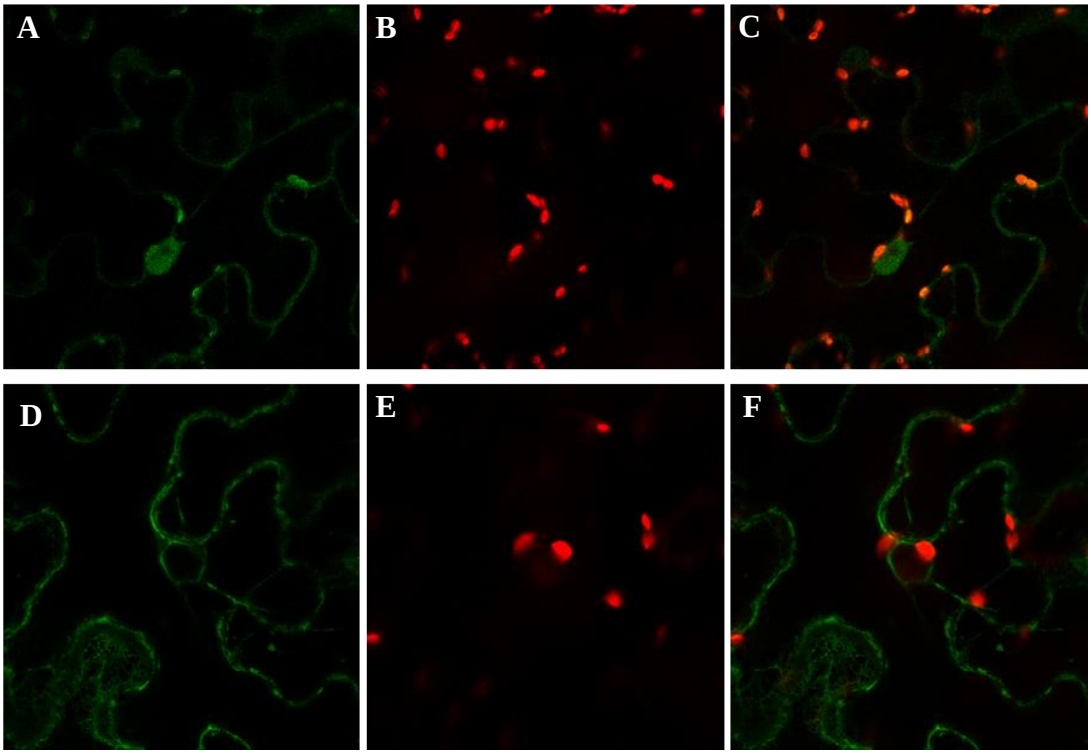


Figure 3.18. Subcellular localization of Pstg11651-GFP inside *N. benthamiana* leaves. (A) Expression of Pstg11651ΔSP-GFP (D) Expression of Pstg11651 (B) and (E) Autofluorescence of the chloroplasts, (C) and (F) Overlaid images. Expression on *N. benthamiana* was achieved using *A. tumefaciens*. The leaves were photographed under confocal microscope (Leica 385 TCS SP5) at 2nd day of the expression. The excitation was at 488 nm and the emission at 495–550 nm (GFP fluorescence) whereas the excitation and emission at far infrared (>800 nm) wavelength to visualize chloroplast autofluorescence.

CHAPTER 4

CONCLUSION

This thesis study is proposed to answer the questions laid upon the key elements of the interaction process in the course of wheat yellow (stripe) disease. Different strategies are benefitted to elucidate the genes involved in the compatible and incompatible interaction of *Pst*. *Pst* is studied extensively by numerous research group all around the world due to extreme importance of the disease epidemiology and the massive economic impact of the targeted host. In similar fashion, *Pst* TR races collected on the fields of Turkey provided by TAGEM (The Central Research Institute for Field Crops) are subjected to de novo transcriptome sequencing using inoculated susceptible and resistant wheat tissues at 10 dpi. The generated data is unique to offer clues about associations between host and pathogen in the manner of previously unstudied time point of 10 dpi, at which the haustoria formation and continuous interface has been established and initial responses are terminated.

The studies are focused on the small-secreted proteins and effector candidates of *Pst* to investigate the proteinaceous factors responsible for establishing, maintaining and ensuring continuous feed from wheat in spite of the presence of the host defenses. Transcriptomics strategy enabled us to catalogue the differentially expressed genes of the disease. Data mining and *in silico* analysis approaches deliver the secretome and effectorome among the differentially expressed genes (DEGs). Moreover, the project created a short cut to pinpoint the most promising candidate list for future effector studies. Additionally, the analysis predicted the possible scenarios that the candidates involve via *in silico* characterization. The comparisons between published reports and database of other *Pst* races and conditions distinguish the common and unique findings in the emphasized disease interval.

This thesis reports the PstDEGs as a consequence of compatible and incompatible interaction of the disease at 10 dpi. However, the main finding is the PstDESSPs, which will serve as an inventory of the effector candidates in future studies. The well-characterized repertoire is an integrative work. For instance, common candidates observed in all compared literature data is more likely to have a function in disease occurrence whereas the unique findings may be specific to Pst TR races and more dispensable in disease progress. Similarly, some of the Avr candidates that are reported in Xia et al., 2017 but dismissed because of the data they compared did not possess. However, they are now back on the track to study. To test the reliability of the PstDESSPs, Unigene17495 (Pstg10917) was investigated in planta studies. The sequence and the prediction of subcellular localization are identical as the forecasts. The finding shows the accuracy of the PstDESSPs and the predictions. Pstg11651 is a candidate identified in Pst78 candidates present in Appendix C and investigated in this study. However, Pstg11651 was thought to be present in PstDESSPs, however the corresponding Unigene is more similar to other Pst78 genes. Nonetheless, the Pst11651 is investigated and identified as chloroplast targeting effector candidate in *N. benthamiana* leaves. Both Pstg10917 and Pst11651 are subjected to cell death suppression assays. Pstg10917 without SP fashioned suppression of cell death symptoms against INF1 elicitor. The finding suggests the localization of Pstg10917 is important, as only Pstg10917 is able to suppress cell death. Pstg11651 was not able to suppress cell death in significant manner.

In this thesis, the assays were also investigated to characterize an effector swiftly. Previously reported effector candidates (PsHa12j12, PstHa15N21 and PStHa12H2) of Yin et al. 2009, was already cloned in pEDV6 vectors. T3SS mediated HR assays on wheat was employed to monitor Avr activity of the candidates. Same expression system used for PTI and cell death suppression assay on both *N. benthamiana* and wheat. None of the effectors reveals significant results for their roles on the host cells. Microarray analysis was conducted on PstTR0997 race inoculated susceptible and resistant wheat cultivars at 24 hpi and 10 dpi. The finding suggest at least more than

200 genes were differentially expressed whereas secreted proteins were not abundant among them. The list of the differentially expressed genes of top hits is reported to serve as a data for comparisons.

Novel findings in this thesis work could be listed as follows;

- Unique expression data collected on 10 dpi of infection, which is the time interval after haustoria is formed and continous interaction between the pathogen and the host is established.
- Novel list of differentially expressed genes during compatible and incompatible interactions using Turkish races.
- In silico characterized catalogue of differentially expressed small secreted proteins (PstDESSPs)
- Integrated outlook on candidate effectors including comparisons with literature work
- Validation of a candidate effector with cell death suppression ability
- Microarray dependent expression profile at 24 hpi and 10 dpi of compatible and incompatible interactions of Turkish Pst races.

These findings will be start point for future effector studies. The well-annotated sets of candidate effectors serve as the most promising candidates for virulence and avirulence attributes of the plant pathogen interaction. For future perspectives, the PstDESSPs list will be studied to elucidate the mechanistic of yellow rust disease at protein level.

In summary, the transcriptome sequencing, data mining, *in silico* characterization, microarray analysis and in planta biological function assays are utilized in an attempt to deliver a promising, reliable and thorough list of candidates of secreted proteins. The candidates will be investigated in the future to decipher their biological functions, virulence and avirulence attributions for the notorious, devastating wheat yellow (stripe) rust disease.

REFERENCES

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research*. https://doi.org/10.1093/nar/25.17.3389

Audic, S., & Claverie, J. M. (1997). The significance of digital gene expression profiles. *Genome Research*, *7*(10), 986–995. https://doi.org/10.1101/gr.7.10.986

Axtell, M. J., & Staskawicz, B. J. (2003). Initiation of RPS2-specified disease resistance in Arabidopsis is coupled to the AvrRpt2-directed elimination of RIN4. *Cell*. https://doi.org/10.1016/S0092-8674(03)00036-9

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., … Noble, W. S. (2009). MEME Suite: Tools for motif discovery and searching. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkp335

Bailey, T. L., & Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. 2nd Int. Conf. on Intelligent Systems for Molecular Biology*. https://doi.org/citeulike-article-id:878292

Bairoch, A., & Apweiler, R. (2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Research*, *28*(1), 45–48. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10592178

Bej, A., Sahoo, B. R., Swain, B., Basu, M., Jayasankar, P., & Samanta, M. (2014). LRRsearch: An asynchronous server-based application for the prediction of leucine-rich repeat motifs and an integrative database of NOD-like receptors. *Computers in Biology and Medicine*, *53*, 164–170. https://doi.org/10.1016/j.compbiomed.2014.07.016

Bendtsen, J. D., Jensen, L. J., Blom, N., Von Heijne, G., & Brunak, S. (2004). Feature-based prediction of non-classical and leaderless protein secretion. *Protein Engineering, Design and Selection*, *17*(4), 349–356. https://doi.org/10.1093/protein/gzh037

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*. https://doi.org/10.1214/aos/1013699998

Boller, T., & Felix, G. (2009). A Renaissance of Elicitors: Perception of Microbe-Associated Molecular Patterns and Danger Signals by Pattern-Recognition Receptors. *Annual Review of Plant Biology*, *60*(1), 379–406. https://doi.org/10.1146/annurev.arplant.57.032905.105346

Bozkurt, T. O., Mcgrann, G. R. D., Maccormack, R., Boyd, L. A., & Akkaya, M. S.

(2010). Cellular and transcriptional responses of wheat during compatible and incompatible race-specific interactions with Puccinia striiformis f. sp. tritici. *Molecular Plant Pathology*, *11*(5), 625–640. https://doi.org/10.1111/j.1364-3703.2010.00633.x

Britton, M., & Cummins, G. B. (1956). The reaction of species of Poa and other grasses to Puccinia striiformis. *Plant Dis Rep.*, *40*, 643–645.

Buchan, D. W. A., Minneci, F., Nugent, T. C. O., Bryson, K., & Jones, D. T. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkt381

Cantu, D., Govindarajulu, M., Kozik, A., Wang, M., Chen, X., Kojima, K. K., … Dubcovsky, J. (2011). Next generation sequencing provides rapid access to the genome of Puccinia striiformis f. sp. tritici, the causal agent of wheat stripe rust. *PLoS ONE*, *6*(8), 4–11. https://doi.org/10.1371/journal.pone.0024230

Cantu, D., Segovia, V., MacLean, D., Bayles, R., Chen, X., Kamoun, S., … Uauy, C. (2013). Genome analyses of the wheat yellow (stripe) rust pathogen Puccinia striiformis f. sp. tritici reveal polymorphic and haustorial expressed secreted proteins as candidate effectors. *BMC Genomics*, *14*(1). https://doi.org/10.1186/1471-2164-14-270

Cao, H., Bowling, S. A., Gordon, A. S., & Dong, X. (1994). Characterization of an Arabidopsis Mutant That Is Nonresponsive to Inducers of Systemic Acquired Resistance. *The Plant Cell*, *6*(11), 1583. https://doi.org/10.2307/3869945

Carbon, S., Ireland, A., Mungall, C. J., Shu, S., Marshall, B., Lewis, S., … Gaudet, P. (2009). AmiGO: Online access to ontology and annotation data. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btn615

Catanzariti, A.-M., Dodds, P. N., Lawrence, G. J., Ayliffe, M. A., & Ellis, J. G. (2006). Haustorially Expressed Secreted Proteins from Flax Rust Are Highly Enriched for Avirulence Elicitors. *The Plant Cell Online*, *18*(1), 243–256. https://doi.org/10.1105/tpc.105.035980

Catanzariti, A.-M., Dodds, P. N., Ve, T., Kobe, B., Ellis, J. G., & Staskawicz, B. J. (2010). The AvrM Effector from Flax Rust Has a Structured C-Terminal Domain and Interacts Directly with the M Resistance Protein. *Molecular Plant-Microbe Interactions*, *23*(1), 49–57. https://doi.org/10.1094/MPMI-23-1-0049

Cesari, S., Bernoux, M., Moncuquet, P., Kroj, T., & Dodds, P. N. (2014). A novel conserved mechanism for plant NLR protein pairs: the â€œintegrated decoyâ€ hypothesis. *Frontiers in Plant Science*, *5*(November), 1–10. https://doi.org/10.3389/fpls.2014.00606

Césari, S., Kanzaki, H., Fujiwara, T., Bernoux, M., Chalvon, V., Kawano, Y., … Kroj, T. (2014). The NB-LRR proteins RGA4 and RGA5 interact functionally and

physically to confer disease resistance. *The EMBO Journal*. https://doi.org/10.15252/embj.201487923

Cesari, S., Thilliez, G., Ribot, C., Chalvon, V., Michel, C., Jauneau, A., … Kroj, T. (2013). The Rice Resistance Protein Pair RGA4/RGA5 Recognizes the Magnaporthe oryzae Effectors AVR-Pia and AVR1-CO39 by Direct Binding. *The Plant Cell*. https://doi.org/10.1105/tpc.112.107201

Chakravarthy, S., Velásquez, A. C., & Martin, G. B. (2009). Assay for Pathogen-Associated Molecular Pattern (PAMP)-Triggered Immunity (PTI) in Plants. *Journal of Visualized Experiments*, (31), 2–5. https://doi.org/10.3791/1442

Chen, X., & Kang, Z. (Eds.). (2017). *Stripe Rust*. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-024-1111-9

Cheng, Y., Wu, K., Yao, J., Li, S., Wang, X., Huang, L., & Kang, Z. (2017). PSTha5a23, a candidate effector from the obligate biotrophic pathogen Puccinia striiformis f. sp. tritici, is involved in plant defense suppression and rust pathogenicity. *Environmental Microbiology*, *19*(5), 1717–1729. https://doi.org/10.1111/1462-2920.13610

Choi, J., Détry, N., Kim, K. T., Asiegbu, F. O., Valkonen, J. P., & Lee, Y. H. (2014). FPoxDB: Fungal peroxidase database for comparative genomics. *BMC Microbiology*. https://doi.org/10.1186/1471-2180-14-117

Choi, J., Park, J., Kim, D., Jung, K., Kang, S., & Lee, Y. H. (2010). Fungal Secretome Database: Integrated platform for annotation of fungal secretomes. *BMC Genomics*. https://doi.org/10.1186/1471-2164-11-105

Cock, P. J. A., & Pritchard, L. (2014). Galaxy as a platform for identifying candidate pathogen effectors. *Methods in Molecular Biology*. https://doi.org/10.1007/978-1-62703-986-4_1

Collier, S. M., & Moffett, P. (2009). NB-LRRs work a "bait and switch" on pathogens. *Trends in Plant Science*. https://doi.org/10.1016/j.tplants.2009.08.001

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., & Robles, M. (2005). Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, *21*(18), 3674–3676. https://doi.org/10.1093/bioinformatics/bti610

Cortázar, A. R., Aransay, A. M., Alfaro, M., Oguiza, J. A., & Lavín, J. L. (2014). SECRETOOL: Integrated secretome analysis tool for fungi. In *Amino Acids*. https://doi.org/10.1007/s00726-013-1649-z

Cuomo, C. A., Bakkeren, G., Khalil, H. B., Panwar, V., Joly, D., Linning, R., … Fellers, J. P. (2017). Comparative Analysis Highlights Variable Genome Content of Wheat Rusts and Divergence of the Mating Loci. *G3: Genes|Genomes|Genetics*, *7*(2), 361–376.

https://doi.org/10.1534/g3.116.032797

Dagvadorj, B., Ozketen, A. C., Andac, A., Duggan, C., Bozkurt, T. O., & Akkaya, M. S. (2017). A Puccinia striiformis f. sp.Tritici secreted protein activates plant immunity at the cell surface. *Scientific Reports*, *7*(1), 1–10. https://doi.org/10.1038/s41598-017-01100-z

Dangl, J. L., Horvath, D. M., & Staskawich, B. J. (2013). Pivoting the Plant Immune System. *Science*, *341*(August), 745–751. https://doi.org/10.1126/science.1236011

Dangl, J. L., & Jones, J. D. G. (2001). Plant pathogens and integrated defence responses to infection. *Nature*. https://doi.org/10.1038/35081161

De Jonge, R., Bolton, M. D., & Thomma, B. P. H. J. (2011). How filamentous pathogens co-opt plants: The ins and outs of fungal effectors. *Current Opinion in Plant Biology*. https://doi.org/10.1016/j.pbi.2011.03.005

De Jonge, R., Van Esse, H. P., Kombrink, A., Shinya, T., Desaki, Y., Bours, R., … Thomma, B. P. H. J. (2010). Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. *Science*. https://doi.org/10.1126/science.1190859

Demirci, Y. E., Inan, C., Günel, A., Maytalman, D., Mert, Z., Baykal, A. T., … Hasançebi, S. (2016). Proteome profiling of the compatible interaction between wheat and stripe rust. *European Journal of Plant Pathology*, *145*(4), 941–962. https://doi.org/10.1007/s10658-016-0882-1

Deslandes, L., Olivier, J., Peeters, N., Feng, D. X., Khounlotham, M., Boucher, C., … Marco, Y. (2003). Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1230660100

Djonović, S., Pozo, M. J., Dangott, L. J., Howell, C. R., & Kenerley, C. M. (2006). Sm1, a Proteinaceous Elicitor Secreted by the Biocontrol Fungus *Trichoderma virens* Induces Plant Defense Responses and Systemic Resistance. *Molecular Plant-Microbe Interactions*. https://doi.org/10.1094/MPMI-19-0838

Dodds, P. N., Lawrence, G. J., Catanzariti, A.-M., Teh, T., Wang, C.-I. A., Ayliffe, M. A., … Ellis, J. G. (2006). Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.0602577103

Dodds, P. N., Lawrence, G. J., Catanzariti, A., Ayliffe, M. A., & Ellis, J. G. (2004). The Melampsora lini AvrL567 Avirulence Genes Are Expressed in Haustoria and Their Products Are Recognized inside Plant Cells. *The Plant Cell Online*, *16*(3),

755–768. https://doi.org/10.1105/tpc.020040

Dodds, P. N., & Rathjen, J. P. (2010). Plant immunity: towards an integrated view of plant–pathogen interactions. *Nature Reviews Genetics*, *11*(8), 539–548. https://doi.org/10.1038/nrg2812

Dou, D., Kale, S. D., Wang, X., Jiang, R. H. Y., Bruce, N. A., Arredondo, F. D., … Tyler, B. M. (2008). RXLR-Mediated Entry of Phytophthora sojae Effector Avr1b into Soybean Cells Does Not Require Pathogen-Encoded Machinery. *Journal of the American Association of Nurse Anesthetists*, *20*(7), 1930–1947. https://doi.org/10.1105/tpc.107.056093

Duplessis, S., Hacquard, S., Delaruelle, C., Tisserant, E., Frey, P., Martin, F., & Kohler, A. (2011). *Melampsora larici-populina* Transcript Profiling During Germination and Timecourse Infection of Poplar Leaves Reveals Dynamic Expression Patterns Associated with Virulence and Biotrophy. *Molecular Plant-Microbe Interactions*, *24*(7), 808–818. https://doi.org/10.1094/MPMI-01-11-0006

Emanuelsson, O., Nielsen, H., Brunak, S., & Von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *Journal of Molecular Biology*. https://doi.org/10.1006/jmbi.2000.3903

Eriksson, J., & Henning, E. (1896). Die Gelreideroste ihre Geschichte und N atur sawie Massregen gegen diesel ben. *Stockholm: P.A. Norstedt Und Soner*, 463.

Evangelisti, E., Gogleva, A., Hainaux, T., Doumane, M., Tulin, F., Quan, C., … Schornack, S. (2017). Time-resolved dual transcriptomics reveal early induced Nicotiana benthamiana root genes and conserved infection-promoting Phytophthora palmivora effectors. *BMC Biology*, *15*(1). https://doi.org/10.1186/s12915-017-0379-1

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., … Punta, M. (2014). Pfam: The protein families database. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkt1223

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., … Bateman, A. (2016). The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkv1344

Fischer, M., & Pleiss, J. (2003). The Lipase Engineering Database: A navigation and analysis tool for protein families. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkg015

Flor, H. H. (1971). Current Status of the Gene-For-Gene Concept. *Annual Review of Phytopathology*. https://doi.org/10.1146/annurev.py.09.090171.001423

Garnica, D. P., Upadhyaya, N. M., Dodds, P. N., & Rathjen, J. P. (2013). Strategies for Wheat Stripe Rust Pathogenicity Identified by Transcriptome Sequencing.

PLoS ONE, *8*(6), e67150. https://doi.org/10.1371/journal.pone.0067150

Gassmann, W., Hinsch, M. E., & Staskawicz, B. J. (1999). The Arabidopsis RPS4 bacterial-resistance gene is a member of the TIR-NBS-LRR family of disease-resistance genes. *Plant Journal*. https://doi.org/10.1046/j.1365-313X.1999.t01-1-00600.x

Germain, H., Joly, D. L., Mireault, C., Plourde, M. B., Letanneur, C., Stewart, D., … Séguin, A. (2018). Infection assays in Arabidopsis reveal candidate effectors from the poplar rust fungus that promote susceptibility to bacteria and oomycete pathogens. *Molecular Plant Pathology*. https://doi.org/10.1111/mpp.12514

Gimenez-Ibanez, S., Hann, D. R., Ntoukakis, V., Petutschnig, E., Lipka, V., & Rathjen, J. P. (2009). AvrPtoB Targets the LysM Receptor Kinase CERK1 to Promote Bacterial Virulence on Plants. *Current Biology*. https://doi.org/10.1016/j.cub.2009.01.054

Godfrey, D., Böhlenius, 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). https://doi.org/10.1186/1471-2164-11-317

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*. https://doi.org/10.1038/nbt.1883

Gu, B., Kale, S. D., Wang, Q., Wang, D., Pan, Q., Cao, H., … Shan, W. (2011). Rust secreted protein Ps87 is conserved in diverse fungal pathogens and contains a RXLR-like motif sufficient for translocation into plant cells. *PLoS ONE*, *6*(11). https://doi.org/10.1371/journal.pone.0027217

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H. Y., Handsaker, R. E., Cano, L. M., … Nusbaum, C. (2009). Genome sequence and analysis of the Irish potato famine pathogen Phytophthora infestans. *Nature*, *461*(7262), 393–398. https://doi.org/10.1038/nature08358

Hacquard, S., Joly, D. L., Lin, Y.-C., Tisserant, E., Feau, N., Delaruelle, C., … Duplessis, S. (2012). A Comprehensive Analysis of Genes Encoding Small Secreted Proteins Identifies Candidate Effectors in *Melampsora larici-populina* (Poplar Leaf Rust). *Molecular Plant-Microbe Interactions*, *25*(3), 279–293. https://doi.org/10.1094/MPMI-09-11-0238

Hall, T. T. A. T. T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Nucleic Acids Symposium Series*. https://doi.org/citeulike-article-id:691774

Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., &

Nakai, K. (2007). WoLF PSORT: Protein localization predictor. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkm259

Huang, X., Chen, X., Coram, T., Wang, M., & Kang, Z. (2011). Gene expression profiling of Puccinia striiformis f. sp. tritici during development reveals a highly dynamic transcriptome. *Journal of Genetics and Genomics*, *38*(8), 357–371. https://doi.org/10.1016/j.jgg.2011.07.004

Hubbard, A., Lewis, C. M., Yoshida, K., Ramirez-Gonzalez, R. H., de Vallavieille-Pope, C., Thomas, J., … Saunders, D. G. O. (2015). Field pathogenomics reveals the emergence of a diverse wheat yellow rust population. *Genome Biology*, *16*(1), 1–15. https://doi.org/10.1186/s13059-015-0590-8

Hückelhoven, R. (2007). Cell Wall–Associated Mechanisms of Disease Resistance and Susceptibility. *Annual Review of Phytopathology*, *45*(1), 101–127. https://doi.org/10.1146/annurev.phyto.45.062806.094325

Iseli, C., Jongeneel, C. V., Bucher, P., & Jongeneel, V. (1999). ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *Proceedings / ...International Conference on Intelligent Systems for Molecular Biology ; ISMB.International Conference on Intelligent Systems for Molecular Biology*. https://doi.org/10.1007/s12576-011-0153-z

Jia, J., Zhao, S., Kong, X., Li, Y., Zhao, G., He, W., … Mao, L. (2013). Aegilops tauschii draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature*. https://doi.org/10.1038/nature12028

Jia, Y., McAdams, S. A., Bryan, G. T., Hershey, H. P., & Valent, B. (2000). Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *EMBO Journal*. https://doi.org/10.1093/emboj/19.15.4004

Jin, Y., Szabo, L. J., & Carson, M. (2010). Century-Old Mystery of *Puccinia striiformis* Life History Solved with the Identification of *Berberis* as an Alternate Host. *Phytopathology*, *100*(5), 432–435. https://doi.org/10.1094/PHYTO-100-5-0432

Jones, D. T., Tress, M., Bryson, K., & Hadley, C. (1999). Successful recognition of protein folds using threading methods biased by sequence similarity and predicted secondary structure. *Proteins: Structure, Function and Genetics*. https://doi.org/10.1002/(SICI)1097-0134(1999)37:3+<104::AID-PROT14>3.0.CO;2-P

Jones, J. D. G., & Dangl, J. L. (2006). The plant immune system. *Nature*, *444*(7117), 323–329. https://doi.org/10.1038/nature05286

Jones, J. D. G., Vance, R. E., & Dangl, J. L. (2016). Intracellular innate immune surveillance devices in plants and animals. *Science*. https://doi.org/10.1126/science.aaf6395

Kaku, H., Nishizawa, Y., Ishii-Minami, N., Akimoto-Tomiyama, C., Dohmae, N., Takio, K., … Shibuya, N. (2006). Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.0508882103

Kale, S. D., Gu, B., Capelluto, D. G. S., Dou, D., Feldman, E., Rumore, A., … Tyler, B. M. (2010). External Lipid PI3P Mediates Entry of Eukaryotic Pathogen Effectors into Plant and Animal Host Cells. *Cell*, *142*(2), 284–295. https://doi.org/10.1016/j.cell.2010.06.008

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., … Yamanishi, Y. (2008). KEGG for linking genomes to life and the environment. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkm882

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, *28*(1), 27–30. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10592173

Kanehisa, M., Sato, Y., & Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *Journal of Molecular Biology*. https://doi.org/10.1016/j.jmb.2015.11.006

Karimi, M., Inzé, D., & Depicker, A. (2002). GATEWAY™ vectors for Agrobacterium-mediated plant transformation. *Trends in Plant Science*. https://doi.org/10.1016/S1360-1385(02)02251-3

Keen, N. T. (1990). Gene-For-Gene Complementarity in Plant-Pathogen Interactions. *Annual Review of Genetics*. https://doi.org/10.1146/annurev.ge.24.120190.002311

Kemen, E., Gardiner, A., Schultz-Larsen, T., Kemen, A. C., Balmuth, A. L., Robert-Seilaniantz, A., … Jones, J. D. G. (2011). Gene gain and loss during evolution of obligate parasitism in the white rust pathogen of Arabidopsis thaliana. *PLoS Biology*, *9*(7). https://doi.org/10.1371/journal.pbio.1001094

Khang, C. H., Berruyer, R., Giraldo, M. C., Kankanala, P., Park, S.-Y., Czymmek, K., … Valent, B. (2010). Translocation of *Magnaporthe oryzae* Effectors into Rice Cells and Their Subsequent Cell-to-Cell Movement. *The Plant Cell*, *22*(4), 1388–1403. https://doi.org/10.1105/tpc.109.069666

Kim, K.-T., Jeon, J., Choi, J., Cheong, K., Song, H., Choi, G., … Lee, Y.-H. (2016). Kingdom-Wide Analysis of Fungal Small Secreted Proteins (SSPs) Reveals their Potential Role in Host Association. *Frontiers in Plant Science*, *7*(February), 1–13. https://doi.org/10.3389/fpls.2016.00186

Kim, M. G., Da Cunha, L., McFall, A. J., Belkhadir, Y., DebRoy, S., Dangl, J. L., & Mackey, D. (2005). Two Pseudomonas syringae type III effectors inhibit RIN4-

regulated basal defense in Arabidopsis. *Cell*. https://doi.org/10.1016/j.cell.2005.03.025

Kobe, B., & Kajava, A. V. (2001). The leucine-rich repeat as a protein recognition motif. *Current Opinion in Structural Biology*, *11*(6), 725–732. https://doi.org/10.1016/S0959-440X(01)00266-4

Krogh, A., Larsson, B., Von Heijne, G., & Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology*. https://doi.org/10.1006/jmbi.2000.4315

Letunic, I., & Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkw290

Letunic, I., & Bork, P. (2018). 20 years of the SMART protein domain annotation resource. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkx922

Letunic, I., Doerks, T., & Bork, P. (2015). SMART: Recent updates, new developments and status in 2015. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gku949

Lévesque, C. A., Brouwer, H., Cano, L., Hamilton, J. P., Holt, C., Huitema, E., … Buell, C. R. (2010). Genome sequence of the necrotrophic plant pathogen Pythium ultimum reveals original pathogenicity mechanisms and effector repertoire. *Genome Biology*, *11*(7). https://doi.org/10.1186/gb-2010-11-7-r73

Ling, P., Wang, M., Chen, X., & Campbell, K. G. (2007). Construction and characterization of a full-length cDNA library for the wheat stripe rust pathogen (Puccinia striiformis f. sp. tritici). *BMC Genomics*, *8*, 1–13. https://doi.org/10.1186/1471-2164-8-145

Liu, C., Pedersen, C., Schultz-Larsen, T., Aguilar, G. B., Madriz-Ordeñana, K., Hovmøller, M. S., & Thordal-Christensen, H. (2016). The stripe rust fungal effector PEC6 suppresses pattern-triggered immunity in a host species-independent manner and interacts with adenosine kinases. *New Phytologist*. https://doi.org/10.1111/nph.14034

Liu, J., Elmore, J. M., Fuglsang, A. T., Palmgren, M. G., Staskawicz, B. J., & Coaker, G. (2009). RIN4 functions with plasma membrane H+-ATPases to regulate stomatal apertures during pathogen attack. *PLoS Biology*. https://doi.org/10.1371/journal.pbio.1000139

Liu, Z., Friesen, T. L., Ling, H., Meinhardt, S. W., Oliver, R. P., Rasmussen, J. B., & Faris, J. D. (2006). The Tsn1 –ToxA interaction in the wheat–Stagonospora nodorum pathosystem parallels that of the wheat–tan spot system. *Genome*, *49*(10), 1265–1273. https://doi.org/10.1139/g06-088

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., & Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkt1178

Lorrain, C., Petre, B., & Duplessis, S. (2018). Show me the way: rust effector targets in heterologous plant systems. *Current Opinion in Microbiology*, *46*, 19–25. https://doi.org/10.1016/j.mib.2018.01.016

Lu, T., Yao, B., & Zhang, C. (2012). DFVF: database of fungal virulence factors. *Database : The Journal of Biological Databases and Curation*. https://doi.org/10.1093/database/bas032

Lum, G., & Min, X. J. (2011). FunSecKB: The Fungal Secretome KnowledgeBase. *Database*, *2011*, 1–10. https://doi.org/10.1093/database/bar001

Mackey, D., Belkhadir, Y., Alonso, J. M., Ecker, J. R., & Dangl, J. L. (2003). Arabidopsis RIN4 is a target of the type III virulence effector AvrRpt2 and modulates RPS2-mediated resistance. *Cell*. https://doi.org/10.1016/S0092-8674(03)00040-0

Mackey, D., Holt, B. F., Wiig, A., & Dangl, J. L. (2002). RIN4 interacts with Pseudomonas syringae type III effector molecules and is required for RPM1-mediated resistance in Arabidopsis. *Cell*. https://doi.org/10.1016/S0092-8674(02)00661-X

Manners, J. G. (1960). Puccinia striiformis Westend. var. Dactylidis var.nov. *Transactions of the British Mycological Society*. https://doi.org/10.1016/S0007-1536(60)80007-1

Manning, V. A., & Ciuffetti, L. M. (2005). Localization of Ptr ToxA Produced by Pyrenophora tritici-repentis Reveals Protein Import into Wheat Mesophyll Cells. *THE PLANT CELL ONLINE*, *17*(11), 3203–3212. https://doi.org/10.1105/tpc.105.035063

Manning, V. a, Hamilton, S. M., Karplus, P. A., & Ciuffetti, L. M. (2008). The Arg-Gly-Asp-containing, solvent-exposed loop of Ptr ToxA is required for internalization. *Molecular Plant-Microbe Interactions : MPMI*, *21*(3), 315–325. https://doi.org/10.1094/MPMI-21-3-0315

Marathe, R., & Dinesh-Kumar, S. P. (2003). Plant defense: One post, multiple guards?! *Molecular Cell*. https://doi.org/10.1016/S1097-2765(03)00072-8

Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., … Bryant, S. H. (2017). CDD/SPARCLE: Functional classification of proteins via subfamily domain architectures. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkw1129

Marchler-Bauer, A., & Bryant, S. H. (2004). CD-Search: Protein domain annotations on the fly. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkh454

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., … Bryant, S. H. (2015). CDD: NCBI's conserved domain database. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gku1221

Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., … Bryant, S. H. (2011). CDD: A Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkq1189

Möller, S., Croning, M. D. R., & Apweiler, R. (2001). Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/17.7.646

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*. https://doi.org/10.1038/nmeth.1226

Mou, Z., Fan, W., Dong, X., & Carolina, N. (2003). nducers of Plant Systemic Acquired Resistance Regulate NPR1 Function through Redox Changes, *113*, 935–944. https://doi.org/10.1016/S0092-8674(03)00429-X

Mucyn, T. S., Wu, A.-J., Balmuth, A. L., Arasteh, J. M., & Rathjen, J. P. (2009). Regulation of tomato Prf by Pto-like protein kinases. *Mol. Plant-Microbe Interact.* https://doi.org/10.1094/MPMI-22-4-0391

Murdoch, C. L., Laemmlen, F. F., & Parvin, P. E. (1973). Response of Kentucky bluegrass and perennial ryegrass cultivars to rust (Puccinia striiformis and P. coronata) in Hawaii. *Plant Dis Rep.*, *57*, 215–216.

Narusaka, M., Shirasu, K., Noutoshi, Y., Kubo, Y., Shiraishi, T., Iwabuchi, M., & Narusaka, Y. (2009). RRS1 and RPS4 provide a dual Resistance-gene system against fungal and bacterial pathogens. *Plant Journal*. https://doi.org/10.1111/j.1365-313X.2009.03949.x

Nickel, W. (2003). The mystery of nonclassical protein secretion: A current view on cargo proteins and potential export routes. *European Journal of Biochemistry*. https://doi.org/10.1046/j.1432-1033.2003.03577.x

Niu, Y. C., Li, Z. Q., & Shang, H. S. (1991). Puccinia striiformis West. f. sp. leymi and f. sp. elymi, two new formae speciales. *Acta Univ Agric Boreali-Occident (Suppl)*, *19*, 58–62.

Nowara, D., Gay, A., Lacomme, C., Shaw, J., Ridout, C., Douchkov, D., … Schweizer, P. (2010). HIGS: Host-Induced Gene Silencing in the Obligate Biotrophic Fungal Pathogen *Blumeria graminis*. *The Plant Cell*, *22*(9), 3130–3141. https://doi.org/10.1105/tpc.110.077040

Ottmann, C., Luberacki, B., Kufner, I., Koch, W., Brunner, F., Weyand, M., … Oecking, C. (2009). A common toxin fold mediates microbial attack and plant

defense. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.0902362106

Panwar, V., McCallum, B., & Bakkeren, G. (2013a). Endogenous silencing of Puccinia triticina pathogenicity genes through in planta-expressed sequences leads to the suppression of rust diseases on wheat. *Plant Journal*, *73*(3), 521–532. https://doi.org/10.1111/tpj.12047

Panwar, V., McCallum, B., & Bakkeren, G. (2013b). Host-induced gene silencing of wheat leaf rust fungus Puccinia triticina pathogenicity genes mediated by the Barley stripe mosaic virus. *Plant Molecular Biology*, *81*(6), 595–608. https://doi.org/10.1007/s11103-013-0022-7

Petersen, T. N., Brunak, S., von Heijne, G., & Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*. https://doi.org/10.1038/nmeth.1701

Petre, B., Lorrain, C., Saunders, D. G. O., Win, J., Sklenar, J., Duplessis, S., & Kamoun, S. (2016). Rust fungal effectors mimic host transit peptides to translocate into chloroplasts. *Cellular Microbiology*. https://doi.org/10.1111/cmi.12530

Petre, B., Saunders, D. G. O., Sklenar, J., Lorrain, C., Krasileva, K. V., Win, J., … Kamoun, S. (2016). Heterologous expression screens in nicotiana benthamiana identify a candidate effector of the wheat yellow rust pathogen that associates with processing bodies. *PLoS ONE*, *11*(2), 1–16. https://doi.org/10.1371/journal.pone.0149035

Petre, B., Saunders, D. G. O., Sklenar, J., Lorrain, C., Win, J., Duplessis, S., & Kamoun, S. (2015). Candidate Effector Proteins of the Rust Pathogen *Melampsora larici-populina* Target Diverse Plant Cell Compartments. *Molecular Plant-Microbe Interactions*. https://doi.org/10.1094/MPMI-01-15-0003-R

Petre, B., Win, J., Menke, F. L. H., & Kamoun, S. (2017). Protein–protein interaction assays with effector–GFP fusions in Nicotiana benthamiana. In *Methods in Molecular Biology*. https://doi.org/10.1007/978-1-4939-7249-4_8

Plett, J. M., Kemppainen, M., Kale, S. D., Kohler, A., Legué, V., Brun, A., … Martin, F. (2011). A secreted effector protein of laccaria bicolor is required for symbiosis development. *Current Biology*, *21*(14), 1197–1203. https://doi.org/10.1016/j.cub.2011.05.033

Pruitt, K. D., Tatusova, T., & Maglott, D. R. (2007). NCBI reference sequences (RefSeq): A curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkl842

Qi, M., Grayczyk, J. P., Seitz, J. M., Lee, Y., Link, T. I., Choi, D., … Whitham, S. A.

(2017). Suppression or Activation of Immune Responses by Predicted Secreted Proteins of the Soybean Rust Pathogen *Phakopsora pachyrhizi*. *Molecular Plant-Microbe Interactions*. https://doi.org/10.1094/MPMI-07-17-0173-FI

Qi, M., Link, T. I., Müller, M., Hirschburger, D., Pudake, R. N., Pedley, K. F., … Whitham, S. A. (2016). A Small Cysteine-Rich Protein from the Asian Soybean Rust Fungus, Phakopsora pachyrhizi, Suppresses Plant Immunity. *PLoS Pathogens*, *12*(9), 1–29. https://doi.org/10.1371/journal.ppat.1005827

Qi, T., Zhu, X., Tan, C., Liu, P., Guo, J., Kang, Z., & Guo, J. (2018). Host-induced gene silencing of an important pathogenicity factor PsCPK1 in Puccinia striiformis f. sp. tritici enhances resistance of wheat to stripe rust. *Plant Biotechnology Journal*, *16*(3), 797–807. https://doi.org/10.1111/pbi.12829

Rafiqi, M., Gan, P. H. P., Ravensdale, M., Lawrence, G. J., Ellis, J. G., Jones, D. A., … Dodds, P. N. (2010). Internalization of Flax Rust Avirulence Proteins into Flax and Tobacco Cells Can Occur in the Absence of the Pathogen. *The Plant Cell*, *22*(6), 2017–2032. https://doi.org/10.1105/tpc.109.072983

Ramachandran, S. R., Yin, C., Kud, J., Tanaka, K., Mahoney, A. K., Xiao, F., & Hulbert, S. H. (2017). Effectors from Wheat Rust Fungi Suppress Multiple Plant Defense Responses. *Phytopathology*, *107*(1), 75–83. https://doi.org/10.1094/PHYTO-02-16-0083-R

Rawlings, N. D., Barrett, A. J., Thomas, P. D., Huang, X., Bateman, A., & Finn, R. D. (2018). The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkx1134

Rawlings, N. D., Waller, M., Barrett, A. J., & Bateman, A. (2014). MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkt953

Ridout, C. J., Skamnioti, P., Porritt, O., Sacristan, S., Jones, J. D. G., Brown, J. K. M., & A. (2006). Multiple Avirulence Paralogues in Cereal Powdery Mildew Fungi May Contribute to Parasite Fitness and Defeat of Plant Resistance. *The Plant Cell Online*, *18*(9), 2402–2414. https://doi.org/10.1105/tpc.106.043307

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*. https://doi.org/10.1371/journal.pone.0029847

Schornack, S., van Damme, M., Bozkurt, T. O., Cano, L. M., Smoker, M., Thines, M., … Huitema, E. (2010). Ancient class of translocated oomycete effectors targets the host nucleus. *Proceedings of the National Academy of Sciences*, *107*(40), 17421–17426. https://doi.org/10.1073/pnas.1008491107

Shabab, M., Shindo, T., Gu, C., Kaschani, F., Pansuriya, T., Chintha, R., … van der Hoorn, R. A. L. (2008). Fungal Effector Protein AVR2 Targets Diversifying Defense-Related Cys Proteases of Tomato. *THE PLANT CELL ONLINE*. https://doi.org/10.1105/tpc.107.056325

Shewry, P. R. (2009). Wheat. *Journal of Experimental Botany*, *60*(6), 1537–1553. https://doi.org/10.1093/jxb/erp058

Shimizu, T., Nakano, T., Takamizawa, D., Desaki, Y., Ishii-Minami, N., Nishizawa, Y., … Shibuya, N. (2010). Two LysM receptor molecules, CEBiP and OsCERK1, cooperatively regulate chitin elicitor signaling in rice. *Plant Journal*. https://doi.org/10.1111/j.1365-313X.2010.04324.x

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., … Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*. https://doi.org/10.1038/msb.2011.75

Sohn, K. H., Lei, R., Nemri, A., & Jones, J. D. G. (2007). The Downy Mildew Effector Proteins ATR1 and ATR13 Promote Disease Susceptibility in Arabidopsis thaliana. *The Plant Cell Online*, *19*(12), 4077–4090. https://doi.org/10.1105/tpc.107.054262

Sonnhammer, E. L. L., von Heijne, G., & Krogh, A. (1998). A hidden Markov model for predicting transmembrane helices in protein sequences. *Proceedings*. https://doi.org/9783223

Sperschneider, J., Catanzariti, A. M., Deboer, K., Petre, B., Gardiner, D. M., Singh, K. B., … Taylor, J. M. (2017). LOCALIZER: Subcellular localization prediction of both plant and effector proteins in the plant cell. *Scientific Reports*. https://doi.org/10.1038/srep44598

Sperschneider, J., Dodds, P. N., Gardiner, D. M., Singh, K. B., & Taylor, J. M. (2018). Improved prediction of fungal effector proteins from secretomes with EffectorP 2.0. *Molecular Plant Pathology*. https://doi.org/10.1111/mpp.12682

Sperschneider, J., Dodds, P. N., Singh, K. B., & Taylor, J. M. (2018). ApoplastP: prediction of effectors and plant proteins in the apoplast using machine learning. *New Phytologist*. https://doi.org/10.1111/nph.14946

Sperschneider, J., Gardiner, D. M., Dodds, P. N., Tini, F., Covarelli, L., Singh, K. B., … Taylor, J. M. (2016). EffectorP: Predicting fungal effector proteins from secretomes using machine learning. *New Phytologist*, *210*(2), 743–761. https://doi.org/10.1111/nph.13794

Stein, K. R., Giardina, B. J., & Chiang, H.-L. (2014). The Non-classical Pathway is the Major Pathway to Secrete Proteins in Saccharomyces cerevisiae. *Clinical & Experimental Pharmacology*, *04*(03). https://doi.org/10.4172/2161-

1459.1000155

Stergiopoulos, I., & de Wit, P. J. G. M. (2009). Fungal Effector Proteins. *Annual Review of Phytopathology*. https://doi.org/10.1146/annurev.phyto.112408.132637

Stotz, H. U., Mitrousia, G. K., de Wit, P. J. G. M., & Fitt, B. D. L. (2014). Effector-triggered defence against apoplastic fungal pathogens. *Trends in Plant Science*. https://doi.org/10.1016/j.tplants.2014.04.009

Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V, … Natale, D. A. (2003). The COG database: an updated version includes eukaryotes. *BMC Bioinformatics*.

Thomas, W. J., Thireault, C. A., Kimbrel, J. A., & Chang, J. H. (2009). Recombineering and stable integration of the Pseudomonas syringae pv. syringae 61 hrp/hrc cluster into the genome of the soil bacterium Pseudomonas fluorescens Pf0-1. *Plant Journal*, *60*(5), 919–928. https://doi.org/10.1111/j.1365-313X.2009.03998.x

Tian, M., Win, J., Song, J., van der Hoorn, R., van der Knaap, E., & Kamoun, S. (2006). A Phytophthora infestans Cystatin-Like Protein Targets a Novel Tomato Papain-Like Apoplastic Protease. *PLANT PHYSIOLOGY*. https://doi.org/10.1104/pp.106.090050

Tollenaar, H. (1967). A comparison of Puccinia striiformis f. sp. poae on bluegrass with P. striiformis f. sp. tritici and f. sp. dactylidis. *Phytopathology*, *57*(4), 418–420.

Tollenaar, H., & Houston, B. R. (1967). A study on the epidemiology of stripe rust, Puccinia striiformis West., in California. *Canadian Journal of Botany*. https://doi.org/10.1139/b67-028

Torto, T. A., Li, S., Styer, A., Huitema, E., Testa, A., Gow, N. A. R., … Kamoun, S. (2003). EST mining and functional expression assays identify extracellular effector proteins from the plant pathogen Phytophthora. *Genome Research*, *13*(7), 1675–1685. https://doi.org/10.1101/gr.910003

Toyota, M., Spencer, D., Sawai-Toyota, S., Jiaqi, W., Zhang, T., Koo, A. J., … Gilroy, S. (2018). Glutamate triggers long-distance, calcium-based plant defense signaling. *Science*, *361*(6407), 1112–1115. https://doi.org/10.1126/science.aat7744

Upadhyaya, N. M., Mago, R., Staskawicz, B. J., Ayliffe, M. A., Ellis, J. G., & Dodds, P. N. (2014). A Bacterial Type III Secretion Assay for Delivery of Fungal Effector Proteins into Wheat. *Molecular Plant-Microbe Interactions*, *27*(3), 255–264. https://doi.org/10.1094/MPMI-07-13-0187-FI

Urban, M., Cuzick, A., Rutherford, K., Irvine, A., Pedro, H., Pant, R., … Hammond-

Kosack, K. E. (2017). PHI-base: A new interface and further additions for the multi-species pathogen-host interactions database. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkw1089

Urban, M., Pant, R., Raghunath, A., Irvine, A. G., Pedro, H., & Hammond-Kosack, K. E. (2015). The Pathogen-Host Interactions database (PHI-base): Additions and future developments. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gku1165

van den Burg, H. A., Harrison, S. J., Joosten, M. H. A. J., Vervoort, J., & de Wit, P. J. G. M. (2006). *Cladosporium fulvum* Avr4 Protects Fungal Cell Walls Against Hydrolysis by Plant Chitinases Accumulating During Infection. *Molecular Plant-Microbe Interactions*. https://doi.org/10.1094/MPMI-19-1420

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. https://doi.org/10.1016/S0968-0004(98)01311-5

van der Hoorn, R. A. L., & Kamoun, S. (2008). From Guard to Decoy: A New Model for Perception of Plant Pathogen Effectors. *THE PLANT CELL ONLINE*. https://doi.org/10.1105/tpc.108.060194

van Esse, H. P., van't Klooster, J. W., Bolton, M. D., Yadeta, K. A., van Baarlen, P., Boeren, S., … Thomma, B. P. H. J. (2008). The Cladosporium fulvum Virulence Protein Avr2 Inhibits Host Proteases Required for Basal Defense. *THE PLANT CELL ONLINE*. https://doi.org/10.1105/tpc.108.059394

Wang, M. N., & Chen, X. M. (2013). First Report of Oregon Grape (Mahonia aquifolium) as an Alternate Host for the Wheat Stripe Rust Pathogen (Puccinia striiformis f. sp. tritici) Under Artificial Inoculation. *Plant Disease*, *97*(6), 839. https://doi.org/10.1094/PDIS-09-12-0864-PDN

Wang, M. N., Wan, A. M., & Chen, X. M. (2015). Barberry Does Not Function as an Alternate Host for Puccinia striiformis f. sp. tritici in the U. S. Pacific Northwest Due to Teliospore Degradation and Barberry Phenology. *Plant Disease*, *99*(11), 1507–1516. https://doi.org/10.1287/msom.1070.0188

Wang, X., Yang, B., Li, K., Kang, Z., Cantu, D., & Dubcovsky, J. (2016). A Conserved *Puccinia striiformis* Protein Interacts with Wheat NPR1 and Reduces Induction of *Pathogenesis - Related* Genes in Response to Pathogens. *Molecular Plant-Microbe Interactions*, *29*(12), 977–989. https://doi.org/10.1094/MPMI-10-16-0207-R

Wellings, C. R. (2011). Global status of stripe rust: A review of historical and current threats. *Euphytica*, *179*(1), 129–141. https://doi.org/10.1007/s10681-011-0360-y

Wellings, C. R., Burdon, J. J., McIntosh, R. A., Wallwork, H., Raman, H., & Murray, G. M. (2000). A new variant of Puccinia striiformis causing stripe rust on barley

and wild Hordeum species in Australia. *Plant Pathology*, *49*(6), 803. https://doi.org/10.1046/j.1365-3059.2000.00506.x

Whisson, S. C., Boevink, P. C., Moleleki, L., Avrova, A. O., Morales, J. G., Gilroy, E. M., … Birch, P. R. J. (2007). A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature*, *450*(7166), 115–118. https://doi.org/10.1038/nature06203

Williams, S. J., Sohn, K. H., Wan, L., Bernoux, M., Sarris, P. F., Segonzac, C., … Jones, J. D. G. (2014). Structural basis for assembly and function of a heterodimeric plant immune receptor. *Science*. https://doi.org/10.1126/science.1247357

Win, J., Kamoun, S., & Jones, A. M. E. (2011). Purification of effector-target protein complexes via transient expression in Nicotiana benthamiana. *Methods in Molecular Biology (Clifton, N.J.)*. https://doi.org/10.1007/978-1-61737-998-7_15

Xia, C., Wan, A., Wang, M., Jiwan, D. A., See, D. R., & Chen, X. (2016). Secreted protein gene derived-single nucleotide polymorphisms (SP-SNPs) reveal population diversity and differentiation of Puccinia striiformis f. sp. tritici in the United States. *Fungal Biology*, *120*(5), 729–744. https://doi.org/10.1016/j.funbio.2016.02.007

Xia, C., Wang, M., Cornejo, O. E., Jiwan, D. A., See, D. R., & Chen, X. (2017). Secretome characterization and correlation analysis reveal putative pathogenicity mechanisms and identify candidate avirulence genes in the wheat stripe rust Fungus Puccinia striiformis f. sp. tritici. *Frontiers in Microbiology*, *8*(DEC). https://doi.org/10.3389/fmicb.2017.02394

Xia, C., Wang, M., Wan, A., Jiwan, D. A., See, D. R., & Chen, X. (2016). Association Analysis of SP-SNPs and Avirulence Genes in Puccinia striiformis f . sp . tritici , the Wheat Stripe Rust Pathogen. *American Journal of Plant Sciences*, (January), 126–137. https://doi.org/10.4236/ajps.2016.71014

Yin, C., Chen, X., Wang, X., Han, Q., Kang, Z., & Hulbert, S. H. (2009). Generation and analysis of expression sequence tags from haustoria of the wheat stripe rust fungus Puccinia striiformis f. sp. Tritici. *BMC Genomics*, *10*, 1–9. https://doi.org/10.1186/1471-2164-10-626

Yin, C., & Hulbert, S. (2011). Prospects for functional analysis of effectors from cereal rust fungi. *Euphytica*, *179*(1), 57–67. https://doi.org/10.1007/s10681-010-0285-x

Yin, C., Jurgenson, J. E., & Hulbert, S. H. (2011). Development of a Host-Induced RNAi System in the Wheat Stripe Rust Fungus *Puccinia striiformis* f. sp. *tritici*. *Molecular Plant-Microbe Interactions*, *24*(5), 554–561. https://doi.org/10.1094/MPMI-10-10-0229

Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., & Xu, Y. (2012). DbCAN: A web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gks479

Yoshida, K., Saitoh, H., Fujisawa, S., Kanzaki, H., Matsumura, H., Yoshida, K., … Terauchia, R. (2009). Association Genetics Reveals Three Novel Avirulence Genes from the Rice Blast Fungal Pathogen Magnaporthe oryzae. *The Plant Cell*, *21*(5), 1573–1591. https://doi.org/10.1105/tpc.109.066324

Yu, C.-S., Chen, Y.-C., Lu, C.-H., & Hwang, J.-K. (2006). Prediction of protein subcellular localization. *Proteins: Structure, Function, and Bioinformatics*, *64*(3), 643–651. https://doi.org/10.1002/prot.21018

Yu, C. S., Cheng, C. W., Su, W. C., Chang, K. C., Huang, S. W., Hwang, J. K., & Lu, C. H. (2014). CELLO2GO: A web server for protein subCELLular lOcalization prediction with functional gene ontology annotation. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0099368

Zadoks, J. C. (1961). Yellow rust on wheat studies in epidemiology and physiologic specialization. *Tijdschrift Voor Plantenziekten*. https://doi.org/10.1007/BF01984044

Zhang, Y., Qu, Z., Zheng, W., Liu, B., Wang, X., Xue, X., … Kang, Z. (2008). Stage-specific gene expression during urediniospore germination in Puccinia striiformis f. sp tritici. *BMC Genomics*, *9*, 1–10. https://doi.org/10.1186/1471-2164-9-203

Zhao, J., Zhuang, H., Zhan, G., Huang, L., & Kang, Z. (2016). Proteomic analysis of Puccina striiformis f. sp. tritici (Pst) during uredospore germination. *European Journal of Plant Pathology*, *144*(1), 121–132. https://doi.org/10.1007/s10658-015-0756-y

Zhao, M., Wang, J., Ji, S., Chen, Z., Xu, J., Tang, C., … Wang, X. (2018). Candidate Effector Pst_8713 Impairs the Plant Immunity and Contributes to Virulence of Puccinia striiformis f. sp. tritici. *Frontiers in Plant Science*, *9*(September), 1–15. https://doi.org/10.3389/fpls.2018.01294

Zhao, Y., Cheng, P., Li, T., & Wang, H. (2018). Proteomic Changes in Urediospores of Puccinia striiformis f . sp . tritici After UV-B Radiation. https://doi.org/10.17957/IJAB/15.0

Zheng, W., Huang, L., Huang, J., Wang, X., Chen, X., Zhao, J., … Kang, Z. (2013). High genome heterozygosity and endemic genetic recombination in the wheat stripe rust fungus. *Nature Communications*, *4*, 1–10. https://doi.org/10.1038/ncomms3673

Zhu, X., Qi, T., Yang, Q., He, F., Tan, C., Ma, W., … Guo, J. (2017). Host-Induced Gene Silencing of the MAPKK Gene *PsFUZ7* Confers Stable Resistance to Wheat Stripe Rust. *Plant Physiology*, *175*(4), 1853–1863.

https://doi.org/10.1104/pp.17.01223

Zipfel, C. (2009). Early molecular events in PAMP-triggered immunity. *Current Opinion in Plant Biology*, *12*(4), 414–420. https://doi.org/10.1016/j.pbi.2009.06.003

Zipfel, C., & Rathjen, J. P. (2008). Plant Immunity: AvrPto Targets the Frontline. *Current Biology*. https://doi.org/10.1016/j.cub.2008.01.016

Zipfel, C., & Robatzek, S. (2010). Pathogen-Associated Molecular Pattern-Triggered Immunity: Veni, Vidi...? *Plant Physiology*, *154*(2), 551–554. https://doi.org/10.1104/pp.110.161547

APPENDICES

A. Plasmid Maps

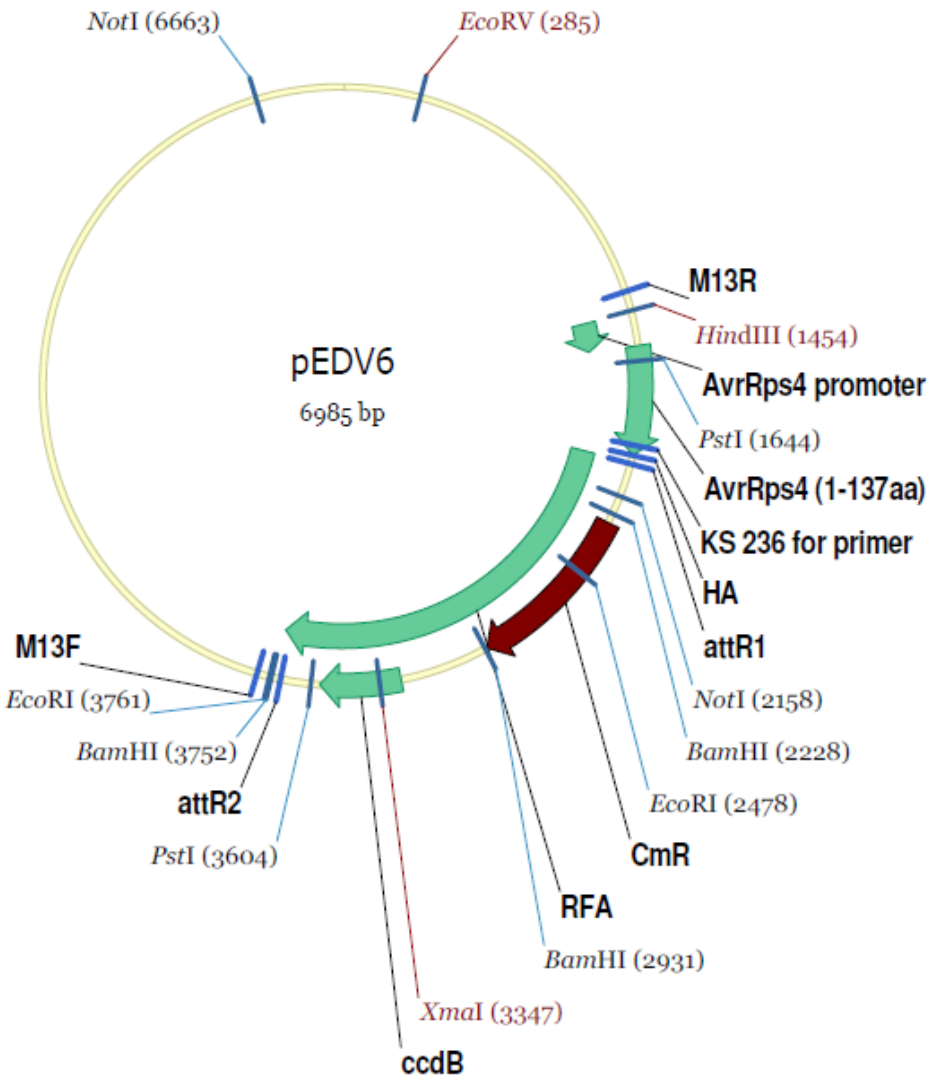


Figure. The plasmid map of pEDV6 vector (Sohn et al., 2007).

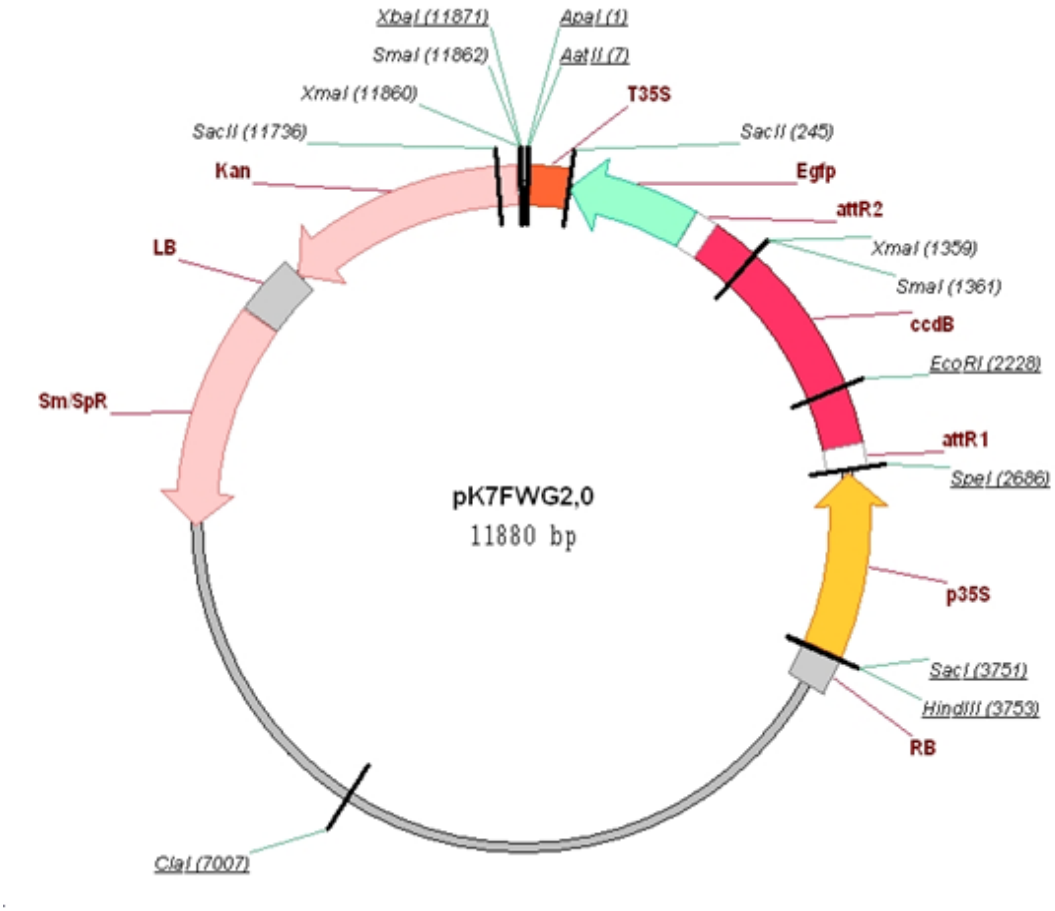


Figure. Plasmid map of pK7FWG2 vector (Karimi et al., 2002).

B. List of The Organelle Targetting Proteins

Table. List of the overlapping genes between transcriptome data and predicted small-secreted proteins from using the pipeline in Section 2.7

Candidate Gene	Length	cTP	mTP	SP	Other	Loc.[a]	ASI*	AYI**
PSTG_06223T0	55	0.907	0.041	0.243	0.167	C	Present	
PSTG_13730T0	113	0.822	0.087	0.061	0.232	C	Present	
PSTG_08405T0	104	0.802	0.144	0.05	0.151	C	Present	
PSTG_13579T0	111	0.678	0.087	0.163	0.283	C	Present	
PSTG_08985T0	70	0.655	0.204	0.021	0.438	C	Present	
PSTG_16102T0	81	0.638	0.053	0.153	0.423	C	Present	
PSTG_16992T0	56	0.579	0.049	0.196	0.572	C	Present	
PSTG_05080T0	100	0.539	0.33	0.011	0.293	C	Present	
PSTG_10917T0	111	0.479	0.185	0.025	0.255	C	Present	Present
PSTG_11111T0	120	0.474	0.217	0.026	0.326	C	Present	Present
PSTG_13486T0	115	0.405	0.258	0.09	0.156	C	Present	
PSTG_03222T0	127	0.365	0.868	0.010	0.015	M	Present	Present
PSTG_16740T0	83	0.026	0.789	0.013	0.401	M	Present	
PSTG_11651T0	104	0.019	0.779	0.020	0.556	M	Present	
PSTG_06238T0	109	0.029	0.755	0.019	0.359	M	Present	
PSTG_15024T0	114	0.04	0.718	0.024	0.430	M	Present	
PSTG_15461T0	100	0.522	0.630	0.021	0.072	M	Present	
PSTG_05065T0	62	0.115	0.613	0.063	0.299	M	Present	
PSTG_05835T0	99	0.382	0.494	0.012	0.301	M	Present	
PSTG_00485T0	109	0.120	0.477	0.033	0.337	M	Present	
PSTG_02632T0	85	0.104	0.448	0.027	0.363	M	Present	
PSTG_15599T0	120	0.046	0.431	0.031	0.006	M	Present	
PSTG_13622T0	115	0.052	0.421	0.169	0.213	M	Present	
PSTG_05064T0	68	0.318	0.380	0.043	0.314	M	Present	

*AvocetS-Pst (10 dpi)**AvocetYR10 (10 dpi)

a The prediction of chloroplast (C) targeting (cTP) or mitochondria (M) targeting (mTP) proteins using TargetP tool (Emmanuelson et al., 2000).

C. Annotation of PstDESSPs

Table. Annotation of PstDESSPs

ID	Description (Blast2GO)	Conserved Domains	Length (no SP)	Cys. Count	Motifs	MEROPS/dbCAN/ LED/fPOXDB
CL1168.Contig1	hyp. prot. PSTG_12786		108	4		
CL1168.Contig2	hyp. prot. PSTG_08840		200	5		
CL1251.Contig1	hyp. prot. PSTG_04504	Tyrosinase superfamily	294	3		
CL1259.Contig1	hyp. prot., variant		94	2		abH13.01 (Bacterial esterase)
CL1259.Contig3	hyp. prot., variant		141	2		
CL1259.Contig4	hyp. prot. PSTG_00428		171	2		
CL1810.Contig1	hyp. prot., variant		110	0		
CL1885.Contig2	---NA---		160	0		DyP-type peroxidase D
CL2026.Contig2	hyp. prot. PSTG_10261		61	1		
CL2255.Contig2	hyp. prot. PSTG_06874, partial		74	0		
CL2673.Contig1	hyp. prot. PSTG_07620		150	1		
CL2966.Contig2	hyp. prot. PSTG_12655		182	2		abH37.01 (*C. antarctica* lipase B like)
CL3094.Contig2	NADH-cytochrome b5 reductase	cyt_b5_reduct_like	235	3		
CL3117.Contig3	hyp. prot. PSTG_07883		58	2		
CL318.Contig1	hyp. prot. PSTG_08979		193	8		
CL318.Contig2	hyp. prot. PSTG_08979		121	6		
CL3516.Contig2	hyp. prot. PSTG_11905		172	2	RXLR=50	
CL3737.Contig2	hyp. prot. PSTG_07428		49	0		
CL3886.Contig1	hyp. prot. PSTG_04394	DUF3455 superfamily	146	2	YXC=48	
CL4066.Contig1	hyp. prot. PSTG_04609		169	12	YXC=165	
CL4121.Contig1	hyp. prot. TRIUR3_32900		57	1		
CL414.Contig2	hyp. prot. PSTG_02885		64	2		
CL4717.Contig1	hyp. prot. PSTG_10402		172	4		
CL4717.Contig2	hyp. prot. PSTG_10402		172	4		
CL4892.Contig2	hyp. prot. PSTG_06134		155	1		
CL499.Contig1	hyp. prot. PSTG_15093	GPI-anchored superfamily	251	3	WXC=86	Hybrid Ascorbate-Cytochrome C peroxidase

CL499.Contig2	hyp. prot. PSTG_15093	GPI-anchored superfamily	251	3	WXC=86	Hybrid Ascorbate-Cytochrome C peroxidase
CL5063.Contig1	hyp. prot. PSTG_02340		43	2		
CL5364.Contig2	hyp. prot. PSTG_01880		205	6		
CL5364.Contig3	hyp. prot. PSTG_01881		205	6		
CL5449.Contig1	hyp. prot. PSTG_11206		121	0		
CL5525.Contig2	and TPR domain	TPR_11 superfamily	207	3	RXLR=189	
CL5553.Contig1	hyp. prot. PSTG_10654		145	3		
CL5553.Contig2	hyp. prot. PSTG_10653		165	2		
CL6430.Contig1	Sm-like ribonucleo		96	6	FXC=94	
CL6446.Contig1	hyp. prot. PSTG_06223		55	2		
CL6446.Contig2	hyp. prot. PSTG_06223		55	2		
CL6541.Contig1	hyp. prot. PSTG_16881		99	0		
CL6640.Contig1	phospholipid-translocating ATPase	Cation_ATPase superfamily	82	0	YXC=11	
CL6786.Contig1	hyp. prot. PSTG_02003		161	3		
CL6816.Contig1	hyp. prot. PSTG_01420		98	8	FXC=113 YXC=48, 102	
CL7100.Contig2	hyp. prot. PSTG_05861		64	2		
CL887.Contig1	hyp. prot. PSTG_16009		138	10	YXC=105, 118	
CL887.Contig2	hyp. prot. PSTG_16009		138	10	YXC=105, 118	
CL913.Contig1	hyp. prot. PSTG_13745		60	0		
Unigene10735	hyp. prot. PSTG_06332		190	2		
Unigene11912	hyp. prot. PSTG_13268		109	10	FXC=98 WXC=52	
Unigene12814	hyp. prot. PSTG_00836		177	2		Hybrid Ascorbate-Cytochrome C peroxidase
Unigene13023	hyp. prot. PSTG_03450		100	10	FXC=99 YXC=38	
Unigene16996	hyp. prot. PSTG_01711	PEBP_euk	272	1		MER0029866
Unigene17495	hyp. prot. PSTG_10917		111	6		
Unigene19081	hyp. prot. PSTG_01012		195	2		
Unigene2221	hyp. prot. PSTG_06810	DUF3129 superfamily	290	9	YXC=81	
Unigene257	hyp. prot. PSTG_06809	DUF3129 superfamily	230	6		
Unigene28265	extracellular conserved serine-rich	GPI-anchored superfamily	195	3		
Unigene28268	hyp. prot. PSTG_11105		77	6	FXC=24	
Unigene28514	hyp. prot. PSTG_11322		128	0		

Unigene28527	hyp. prot. PSTG_13630		146	0		abH04.02 (Acinetobacter esterases)
Unigene28585	hyp. prot., variant	CFEM	166	9	FXC=55	
Unigene28734	subtilisin protease	Peptidases_S8_S53 superfamily	138	3		abH08.13 (soluble esterases /lipases /peptidases)
Unigene28760	Non-Catalytic module family EXPN	DPBB_1 superfamily	112	6		
Unigene28808	---NA---		160	4	WXC=91	
Unigene29022	peptidyl-prolyl cis-trans isomerase B	cyclophilin superfamily	197	1	FXC=154	
Unigene29087	hyp. prot. PSTG_07364	Cupredoxin superfamily	170	3	YXC=128	
Unigene29134	hyp. prot. PSTG_17553	Ribophorin_II superfamily	117	0		
Unigene29157	hyp. prot. PSTG_16911		208	0		
Unigene29179	hyp. prot. PSTG_04274		82	6	FXC=78 YXC=31	
Unigene29215	hyp. prot. PSTG_11332	ML superfamily	168	6		
Unigene30526	P-loop containing nucleoside triphosphate hydrolase	HA2 superfamily	81	0		
Unigene30571	hyp. prot. PGTG_21113	Yos1	55	0		
Unigene30573	hyp. prot. PSTG_06099	BP28CT superfamily	94	0		
Unigene30755	probable rcd1 involved in sexual development		55	4		
Unigene30776	hyp. prot. PSTG_08831		108	0		
Unigene30809	hyp. prot. PSTG_03083		46	0		
Unigene30875	hyp. prot. PSTG_03524	PDI_a_family	184	4		
Unigene30934	hyp. prot. PSTG_02921		141	2		
Unigene31019	hyp. prot. PSTG_07038		115	0		
Unigene31042	hyp. prot. PSTG_01876		55	4		
Unigene31064	hyp. prot. PSTG_02655		109	0		Rbohs
Unigene31121	hyp. prot. PSTG_13576		39	0	YXC=3	
Unigene31124	ATP-dependent peptidase	TIP49 superfamily	63	1		
Unigene31157	hyp. prot. PGTG_15897	Rot1 superfamily	205	2		
Unigene31162	hyp. prot. PSTG_14206		20	0		
Unigene31232	hyp. prot. PTTG_07345		87	8		
Unigene31238	hyp. prot. PGTG_11235		70	0		
Unigene31294	hyp. prot. PSTG_00994		254	10		
Unigene31299	hyp. prot. PSTG_01843		92	0		
Unigene31341	hyp. prot. PSTG_14884	Neuromodulin_N superfamily	237	1	RXLR=144	Fungi-Bacteria glutathione peroxidase

Unigene31480	---NA---		29	2	YXC=30	
Unigene31503	hyp. prot. PSTG_14184		53	0		
Unigene31618	hyp. prot. PSTG_17090	RNase_E_G superfamily	151	2		
Unigene31734	1,3-beta-glucan synthase component FKS1	Glucan_synthase superfamily	83	2		
Unigene31812	hyp. prot. PSTG_13576		118	0	YXC=3	
Unigene31899	hyp. prot. PSTG_17264		94	0	RXLR=77	
Unigene31902	hyp. prot. PSTG_07799	vWFA superfamily	48	4		
Unigene31932	hyp. prot. PSTG_08524		100	0	RXLR=95	
Unigene31938	copper zinc superoxide dismutase	Sod_Cu	165	4		
Unigene3203	hyp. prot. PSTG_11850		136	6		
Unigene32092	glucose-regulated variant	NBD_sugar-kinase_HSP70_actin superfamily	129	1	YXC=49	
Unigene32286	hyp. prot. PSTG_01135		106	6	FXC=33 YXC=78	
Unigene32343	---NA---		51	0		
Unigene32502	hyp. prot. PSTG_08409		102	2	FXC=124	
Unigene32923	hyp. prot. PSTG_07376		34	0		
Unigene33226	hyp. prot. PSTG_13913		72	6	YXC=29	
Unigene33242	hyp. prot. PSTG_16262		94	6		
Unigene33243	6-phosphofructokinase subunit beta	PFK superfamily	82	2		
Unigene33278	hyp. prot. PSTG_07099		65	0		
Unigene33551	hyp. prot. PSTG_05514		200	2		
Unigene33596	hyp. prot. PSTG_13486		115	8	YXC=58, 124	
Unigene33762	---NA---		21	1		
Unigene33801	hyp. prot. PSTG_05079		92	6	FXC=29 YXC=67	
Unigene33803	hyp. prot. PSTG_13581		50	0		
Unigene33834	hyp. prot. PSTG_01770		126	0		
Unigene33938	rust transferred		158	6		
Unigene34056	hyp. prot. PSTG_13667		66	5	YXC=57	
Unigene34113	glycoside hydrolase family 63	Glyco_hydro_63N	191	1		
Unigene34381	---NA---		51	5		
Unigene34390	U6 snRNA-associated Sm LSm3		66	1	FXC=11	
Unigene34485	pyridoxine synthesis PDX2	GAT_1 superfamily	126	0		MER0066916

Unigene34520	FK506-binding 2	FKBP_C	131	2		
Unigene34564	glycoside hydrolase family 63		162	2		
Unigene34612	glycosyltransferase family 2	Cyt-b5	183	1		
Unigene34643	hyp. prot. PSTG_01892		57	1		
Unigene34687	hyp. prot. PSTG_09177		95	2		
Unigene34775	hyp. prot. PSTG_15691	LanC_like superfamily	211	3		GH76.hmm
Unigene34787	hyp. prot. PSTG_14510		112	3	FXC=59, 105	
Unigene35064	homoaconitate hydratase	AcnA superfamily	154	0		
Unigene35089	hyp. prot. PSTG_06253	PG-PI_TP	171	4		
Unigene35403	hyp. prot. PSTG_17053		84	6	FXC=29 YXC=67, 100 WXC=77	
Unigene35404	hyp. prot. PSTG_17053		83	6	FXC=29 YXC=67 WXC=77	
Unigene35538	alpha beta-hydrolase	Abhydrolase superfamily	142	3		abH19.01 (Palmitoyl-protein thioesterase 1 like)
Unigene35601	hyp. prot. PSTG_02688		43	1		
Unigene35724	hyp. prot. PSTG_03362	Glyco_hydro_2 superfamily	234	1		abH32.01 (Xylanase Z esterase domain)
Unigene35735	Sec7 domain-containing		135	0		
Unigene35965	hyp. prot. PSTG_09603		28	0		
Unigene35969	hyp. prot., variant 3		117	1		
Unigene36059	hyp. prot. PSTG_03254		46	3		
Unigene36089	hyp. prot. PSTG_07659		64	0		
Unigene36187	hyp. prot. PSTG_00676		71	2		
Unigene36196	hyp. prot. PSTG_14086		284	10		
Unigene36211	hyp. prot. PSTG_11460		120	0		
Unigene36268	hyp. prot. PSTG_16164		240	2		Hybrid Ascorbate-Cytochrome C peroxidase
Unigene36307	hyp. prot. PSTG_06305		163	7		
Unigene36315	COPII-coated vesicle [Rhodotorula toruloides NP11]	EMP24_GP25L	162	2		
Unigene36335	family 71 glycoside hydrolase	Glyco_hydro_71 superfamily	239	2		GH71.hmm
Unigene36336	hyp. prot. PSTG_06808		164	0		
Unigene36354	hyp. prot. PSTG_08969		196	8		
Unigene36379	hyp. prot. PSTG_15910	mTERF superfamily	150	0		
Unigene36535	hyp. prot. PSTG_02031		36	0		
Unigene36568	hyp. prot. PSTG_03661		144	3	WXC=34	

Unigene36846	glyco endopeptidase KAE1	Peptidase_M22 superfamily	113	2		
Unigene36850	hyp. prot. PSTG_01696		66	0		
Unigene36901	hyp. prot. PSTG_03879, partial		49	6		
Unigene37161	hyp. prot. PSTG_00209		117	0	WXC=34	
Unigene37172	hyp. prot. PSTG_05297	YoaJ superfamily	266	6		CBM63.hmm
Unigene37241	hyp. prot. PSTG_11923		76	1		
Unigene37294	hyp. prot., variant		66	0		
Unigene37534	hyp. prot. PSTG_02897		46	0		
Unigene37828	hyp. prot. PSTG_01576	DUF3605 superfamily	141	0		
Unigene37852	hyp. prot. PSTG_00777		37	0		
Unigene37856	STE STE7 MEK1 kinase	PKc_like superfamily	157	2	YXC=126	
Unigene37915	hyp. prot. PSTG_09282		157	12	FXC=26, 108 YXC=58, 138	
Unigene37966	hyp. prot. PSTG_13342	SCP superfamily	139	1		
Unigene37988	hyp. prot. PSTG_16474	DNA_pol3_gamma3 superfamily	153	0		
Unigene38087	hyp. prot. PSTG_09694		121	3		
Unigene38169	hyp. prot. PSTG_05869	DUF2370 superfamily	95	0		
Unigene38264	L-ascorbate oxidase	SufI superfamily	218	5	FXC=235	
Unigene38311	hyp. prot. PSTG_04571		236	0		
Unigene38374	family 5 glycoside hydrolase family 13 glycosyltransferase	AmyAc_family superfamily	219	2	YXC=124	GH13.hmm/abH32.02 (Xylanase Y esterase domain)
Unigene38507	CMGC CLK kinase	PKc_like superfamily	130	1	FXC=5	abH32.01 (Xylanase Z esterase domain)
Unigene38512	hyp. prot. PSTG_14565		142	12	FXC=108	
Unigene38530	hyp. prot. PTTG_00528		206	6		
Unigene38531	hyp. prot. PSTG_19591, partial		61	5	WXC=63	
Unigene38573	ribonuclease T2	Ribonuclease_T2	187	8		
Unigene38612	hyp. prot. PSTG_14728	GPI-anchored / ALDH-SF superfamily	131	2	WXC=85	
Unigene38637	adenosylmethionine decarboxylase	SAM_decarbox superfamily	110	1		
Unigene38681	hyp. prot. PSTG_09146	5_nucleotid_C superfamily	245	6		Hybrid Ascorbate-Cytochrome C peroxidase
Unigene38718	hyp. prot. PSTG_01423		195	14	FXC=22, 139, 197, 208 YXC=58	
Unigene38721	hyp. prot. PSTG_04258		67	1		

Unigene38765	hyp. prot. PSTG_00149		221	2		
Unigene38778	hyp. prot. PSTG_13886		197	9		
Unigene38807	hyp. prot. PSTG_14378		191	10	FXC=125	
Unigene38824	transcription elongation factor SPT4 [Rhodotorula toruloides NP11]		94	5		
Unigene38928	carboxypeptidase C	Peptidase_S10 superfamily	202	3		MER0003541/abH34.02 (Serine carboxypeptidase II like)
Unigene39100	hyp. prot. PSTG_14695		132	6	WXC=46	
Unigene39119	hyp. prot. PSTG_14557	Ribosomal_L6 superfamily	265	0	RXLR=63	
Unigene39120	---NA---		86	2		DyP-type peroxidase D
Unigene39125	hyp. prot. PSTG_04309	pepsin_retropepsin_like superfamily	130	2		MER0000941
Unigene39129	hyp. prot. PSTG_12571		51	6	FXC=29, 52	
Unigene39165	hyp. prot. PSTG_02113		90	2		
Unigene39169	disulfide-isomerase A6 [Rhodotorula toruloides NP11]	Thioredoxin_like superfamily / ERp29	241	4		
Unigene39212	hyp. prot. PSTG_13167		204	6		abH08.13 (soluble esterases / lipases / peptidases)
Unigene39228	hyp. prot. PSTG_00629		113	2		
Unigene39262	hyp. prot. PSTG_03370		126	0	RXLR=141	
Unigene39288	hyp. prot. PSTG_14933	CFEM	72	8	FXC=51	
Unigene39340	family 5 glycoside hydrolase		67	1		
Unigene39371	hyp. prot. PSTG_06808	DnaJ superfamily	64	0		
Unigene39373	---NA---		16	0		
Unigene39381	ubiquitin-conjugating enzyme E2 J2		41	1		
Unigene39382	hyp. prot. PSTG_11331		111	1		
Unigene39386	RHTO0S09e06304g1_1 [Rhodotorula toruloides]	DUF1748 superfamily	69	0		
Unigene39390	hyp. prot. PSTG_16882		107	0		
Unigene39396	hyp. prot. PSTG_08709		57	6	YXC=38, 46	
Unigene39468	hyp. prot. PSTG_00445		36	0		
Unigene39525	---NA---		85	2		
Unigene39540	hyp. prot. PSTG_17409		46	5	FXC=66	
Unigene39555	hyp. prot. PSTG_07423		42	4		

Unigene39647	hyp. prot. PSTG_15461		100	4		
Unigene39691	hyp. prot. PSTG_00906		69	0		
Unigene39857	hyp. prot. PSTG_08566		183	1		
Unigene39905	hyp. prot. PSTG_15857		85	1		
Unigene39933	hyp. prot. PSTG_03634		123	2		
Unigene40000	hyp. prot. PSTG_05654		64	1		
Unigene40066	hyp. prot. PSTG_08556		209	3	RXLR=129	abH08.07 (soluble epoxide hydrolases (beta6))
Unigene40134	hyp. prot. PSTG_10610		43	0		
Unigene40198	hyp. prot. PSTG_14414		92	1		
Unigene40532	hyp. prot. PSTG_13861		110	0	RXLR=55	
Unigene40600	hyp. prot. PSTG_04328	DUF775 superfamily	87	1		
Unigene40839	hyp. prot. PSTG_09471		112	1		
Unigene40952	cytochrome c oxidase assembly [Rhodotorula toruloides NP11]	PET117	84	2		
Unigene41224	hyp. prot. PTTG_27311		85	6		
Unigene41262	hyp. prot. PSTG_04871		87	0		
Unigene41839	hyp. prot. PSTG_09266		87	2		
Unigene4423	hyp. prot. PSTG_06021		175	2		
Unigene477	hyp. prot. PSTG_15335		101	0		
Unigene5754	hyp. prot. PSTG_08755	DPBB_1 superfamily	231	5		
Unigene7586	hyp. prot. PSTG_14207		218	3	FXC=10	
Unigene7732	uncharacterized protein LOC109745535		185	7	FXC=158 RXLR=56	
Unigene7930	hyp. prot. PSTG_02173		109	4		
Unigene9158	hyp. prot. PSTG_01739		93	0		

D. Characterization of PstDESSPs

Table. Differential expression analysis and subcellular localization prediction.

	Differential Expression Analysis			**Subcellular Localization Prediction**			
ID	Avocet-S-*Pst*/ Avocet-YR10-*Pst*	P-value	FDR	TargetP	Localizer	WolfPSORT	ApoplastP
CL1168.Contig1	Up	1,01E-33	6,10E-24	C		N	Non-apoplastic
CL1168.Contig2	Up	1,73E-156	3,91E-146	C	C,M	C	Non-apoplastic
CL1251.Contig1	Up	1,02E-01	1,01E+08	C	M	N	Non-apoplastic
CL1259.Contig1	Up	2,63E-05	3,74E+04	C		C	Apoplastic
CL1259.Contig3	Up	1,81E-02	2,04E+07	_		N	Apoplastic
CL1259.Contig4	Up	7,55E-59	7,39E-48	_		N	Apoplastic
CL1810.Contig1	Up	0	0	_		N	Apoplastic
CL1885.Contig2	Down	4,42E-03	4,71E+07	_		N	Apoplastic
CL2026.Contig2	Up	9,13E-07	1,36E+04	M	N	N	Non-apoplastic
CL2255.Contig2	Up	1,52E-21	6,23E-12	_		cyto	Apoplastic
CL2673.Contig1	Up	6,87E-13	1,79E-02	_		golg	Non-apoplastic
CL2966.Contig2	Up	1,36E-10	3,09E+00	M	M,C	N	Non-apoplastic
CL3094.Contig2	Up	8,15E-04	9,63E+06	M	M	M	Non-apoplastic
CL3117.Contig3	Up	5,38E-06	8,22E+04	_		C	Non-apoplastic
CL318.Contig1	Up	9,68E-12	2,36E-01	_	N	N	Non-apoplastic
CL318.Contig2	Up	2,99E+00	0.000188	_		C	Non-apoplastic
CL3516.Contig2	Up	3,82E-152	8,30E-141	C	C,M	C	Apoplastic
CL3737.Contig2	Up	6,29E-04	7,58E+06	_		cyto	Non-apoplastic
CL3886.Contig1	Up	2,04E-263	9,11E-252	M		C	Apoplastic
CL4066.Contig1	Up	1,14E-46	9,19E-36	C		C	Apoplastic
CL4121.Contig1	Down	8,54E+00	0.000482	M		nucl_plas	Non-apoplastic
CL414.Contig2	Up	6,16E-03	6,37E+07	C		M	Non-apoplastic
CL4717.Contig1	Up	7,58E-05	1,02E+06	_		C	Non-apoplastic
CL4717.Contig2	Up	1,76E+01	0.000117	_		chlo	Non-apoplastic
CL4892.Contig2	Up	2,07E-35	1,28E-24	_		mito	Non-apoplastic
CL499.Contig1	Up	5,64E-29	2,92E-18	_		mito	Apoplastic
CL499.Contig2	Up	1,55E-04	2,26E+04	_		mito	Apoplastic
CL5063.Contig1	Up	2,35E+00	0.000150	_		nucl	Non-apoplastic
CL5364.Contig2	Up	1,15E-28	6,13E-19	_		E.R.	Non-apoplastic
CL5364.Contig3	Up	1,83E-03	2,36E+06	_		nucl	Apoplastic
CL5449.Contig1	Up	0	0	_		cyto	Apoplastic
CL5525.Contig2	Up	5,54E-09	1,10E+02	_		nucl	Non-apoplastic
CL5553.Contig1	Up	1,24E-09	2,80E+00	_	N	nucl	Non-apoplastic

CL5553.Contig2	Up	3,77E-163	8,89E-152	_		nucl	Non-apoplastic
CL6430.Contig1	Up	1,55E-04	2,26E+05	C	C	nucl	Non-apoplastic
CL6446.Contig1	Up	5,52E-63	5,72E-52	C		chlo	Apoplastic
CL6446.Contig2	Up	1,15E-67	1,30E-57	C		chlo	Apoplastic
CL6541.Contig1	Up	8,52E-17	2,74E-06	_	N	cyto	Non-apoplastic
CL6640.Contig1	Up	3,60E-01	2,73E+09	_		nucl	Non-apoplastic
CL6786.Contig1	Up	1,53E-21	6,29E-12	M	C	nucl	Non-apoplastic
CL6816.Contig1	Up	3,70E-12	8,68E-01	_		chlo	Apoplastic
CL7100.Contig2	Up	3,71E-11	8,70E-01	M		cyto	Apoplastic
CL887.Contig1	Up	2,32E-05	3,34E+05	_		chlo	Apoplastic
CL887.Contig2	Up	2,52E-29	1,28E-17	_		chlo	Apoplastic
CL913.Contig1	Up	1,04E+00	7,21E+09	C		nucl	Apoplastic
Unigene10735	Up	0	0	C		nucl	Apoplastic
Unigene11912	Up	2,44E-179	6,37E-168	_		nucl	Apoplastic
Unigene12814	Up	1,34E-180	3,57E-170	_		extr	Apoplastic
Unigene13023	Up	0	0	_		chlo	Apoplastic
Unigene16996	Up	3,19E-06	4,96E+04	_		golg	Apoplastic
Unigene17495	Up	0	0	C	C	mito	Apoplastic
Unigene19081	Up	0	0	_		nucl	Apoplastic
Unigene2221	Up	0	0	_		chlo	Non-apoplastic
Unigene257	Up	1,75E-129	3,27E-120	_		chlo	Apoplastic
Unigene28265	Up	0	0	_		extr	Apoplastic
Unigene28268	Up	1,40E-06	2,05E+04	_		cyto	Apoplastic
Unigene28514	Up	2,31E-11	5,13E+00	_		nucl	Non-apoplastic
Unigene28527	Up	1,87E-05	2,98E+04	C	C	nucl	Non-apoplastic
Unigene28585	Up	0	0	_		chlo	Apoplastic
Unigene28734	Up	9,96E-15	2,87E-04	_		nucl	Apoplastic
Unigene28760	Up	1,14E-46	9,17E-36	_		chlo	Apoplastic
Unigene28808	Up	0.0001767142	0.000902	_	C	nucl	Non-apoplastic
Unigene29022	Up	7,92E-82	1,04E-70	_	C,M	cyto	Non-apoplastic
Unigene29087	Up	4,91E-78	6,17E-67	_		mito	Apoplastic
Unigene29134	Up	2,68E-25	1,21E-14	_		cyto	Apoplastic
Unigene29157	Up	3,90E-118	6,70E-107	_		nucl	Non-apoplastic
Unigene29179	Up	0	0	_		chlo	Apoplastic
Unigene29215	Up	4,53E-57	4,36E-46	_		chlo	Apoplastic
Unigene30526	Up	1,04E+00	7,25E+09	_		nucl	Non-apoplastic
Unigene30571	Up	0.0001461314	0.000758	_		cyto	Non-apoplastic
Unigene30573	Up	6,11E-01	4,44E+08	S		nucl	Non-apoplastic
Unigene30755	Up	0.0001461314	0.000766	C		extr	Non-apoplastic

Unigene30776	Up	6,11E-01	4,43E+09	_		nucl	Non-apoplastic
Unigene30809	Up	5,07E+00	0.000297	_		chlo	Apoplastic
Unigene30875	Up	1,78E-43	1,36E-33	_		nucl	Non-apoplastic
Unigene30934	Up	5,22E-03	5,48E+07	_		nucl	Apoplastic
Unigene31019	Up	6,11E-01	4,40E+09	C		chlo	Non-apoplastic
Unigene31042	Up	8,03E-11	1,84E+00	S		cyto	Apoplastic
Unigene31064	Up	1,64E-11	3,93E-01	C		nucl	Apoplastic
Unigene31121	Up	5,38E-06	8,20E+04	_		nucl	Non-apoplastic
Unigene31124	Up	1,76E+01	0.000116	_		extr	Non-apoplastic
Unigene31157	Up	2,10E-19	7,28E-08	_		nucl	Non-apoplastic
Unigene31162	Up	3,82E-08	6,58E+03	_		-	Non-apoplastic
Unigene31232	Up	2,52E-73	3,01E-62	_		chlo	Apoplastic
Unigene31238	Up	2,45E-17	7,77E-06	M		chlo	Non-apoplastic
Unigene31294	Up	1,98E-11	5,02E-02	_	N	nucl	Non-apoplastic
Unigene31299	Up	2,63E-05	3,77E+05	_		nucl	Non-apoplastic
Unigene31341	Up	3,08E-65	3,29E-54	_	N	nucl	Non-apoplastic
Unigene31480	Up	2,18E-04	2,78E+06	_		-	Apoplastic
Unigene31503	Up	5,07E+00	0.000296	_		chlo	Non-apoplastic
Unigene31618	Up	1,31E-32	7,70E-23	S	M,C	chlo	Non-apoplastic
Unigene31734	Up	5,07E+00	0.000298	_		chlo	Non-apoplastic
Unigene31812	Up	1,07E-02	1,24E+07	_		nucl	Non-apoplastic
Unigene31899	Up	1,07E-02	1,24E+07	C		nucl	Non-apoplastic
Unigene31902	Up	1,29E-94	1,90E-84	_		chlo	Apoplastic
Unigene31932	Up	1,76E+01	0.000116	C		cyto_nucl	Non-apoplastic
Unigene31938	Up	1,43E-177	3,68E-167	_		nucl	Apoplastic
Unigene3203	Up	0.0001428924	0.000762	_		cysk	Apoplastic
Unigene32092	Up	1,17E-103	1,85E-92	_	N	cyto	Non-apoplastic
Unigene32286	Up	5,51E-270	2,57E-258	_		chlo	Apoplastic
Unigene32343	Up	8,61E+00	0.000478	M		cyto_nucl	Non-apoplastic
Unigene32502	Up	0	0	C		nucl	Apoplastic
Unigene32923	Up	3,37E-104	5,31E-92	_		nucl	Apoplastic
Unigene33226	Up	6,95E-117	1,18E-105	M		nucl	Apoplastic
Unigene33242	Up	1,16E-65	1,27E-55	C		chlo	Apoplastic
Unigene33243	Up	7,60E-02	6,53E+07	_		cyto	Non-apoplastic
Unigene33278	Up	7,72E-25	3,43E-14	M		mito	Apoplastic
Unigene33551	Up	4,72E-25	2,11E-15	C		chlo	Apoplastic
Unigene33596	Up	5,02E-17	1,63E-06	C	C	chlo	Non-apoplastic
Unigene33762	Up	1,17E-11	3,00E-02	_		-	Non-apoplastic
Unigene33801	Up	0	0	_		chlo	Apoplastic

Unigene33803	Up	6,48E-07	1,10E+04	C		nucl	Apoplastic
Unigene33834	Up	9,96E-13	2,57E-02	_		chlo	Non-apoplastic
Unigene33938	Up	2,18E-04	2,79E+06	_		chlo	Apoplastic
Unigene34056	Up	8,61E+00	0.000480	M		extr	Non-apoplastic
Unigene34113	Up	5,87E-15	1,72E-04	_		nucl	Non-apoplastic
Unigene34381	Up	3,63E-38	2,39E-27	_		chlo	Apoplastic
Unigene34390	Up	1,33E-24	6,09E-15	M		nucl	Non-apoplastic
Unigene34485	Up	2,07E-01	1,67E+05	_		cyto_nucl	Non-apoplastic
Unigene34520	Up	1,29E-127	2,38E-117	_	N	mito	Non-apoplastic
Unigene34564	Up	4,63E-46	3,57E-34	_		mito	Non-apoplastic
Unigene34612	Up	2,55E-02	2,39E+08	_		cyto	Apoplastic
Unigene34643	Up	2,01E-32	1,15E-21	_		mito	Non-apoplastic
Unigene34687	Up	1,76E+01	0.000115	C		cyto	Non-apoplastic
Unigene34775	Up	4,71E-99	7,24E-88	_		nucl	Non-apoplastic
Unigene34787	Up	1,55E-04	2,26E+05	C		nucl	Non-apoplastic
Unigene35064	Up	1,25E+00	1,05E+09	C	N	cyto_nucl	Apoplastic
Unigene35089	Up	1,59E-118	2,76E-107	_		cyto	Non-apoplastic
Unigene35403	Up	6,30E-26	2,94E-15	_		cyto	Apoplastic
Unigene35404	Up	8,48E-172	2,10E-160	_		chlo	Apoplastic
Unigene35538	Up	6,68E-04	7,95E+06	_		extr	Non-apoplastic
Unigene35601	Up	1,76E+01	0.000116	M		cyto	Non-apoplastic
Unigene35724	Up	2,45E-17	7,70E-06	C	C	chlo	Non-apoplastic
Unigene35735	Up	2,55E-02	2,41E+08	_		nucl	Non-apoplastic
Unigene35965	Up	2,55E-02	2,39E+07	_		-	Non-apoplastic
Unigene35969	Up	2,79E-02	2,59E+08	_		cyto	Non-apoplastic
Unigene36059	Up	4,34E-02	3,95E+08	_		mito	Non-apoplastic
Unigene36089	Up	4,34E-02	3,91E+08	C		cyto_nucl	Apoplastic
Unigene36187	Up	6,87E-13	1,79E-02	_		cyto	Apoplastic
Unigene36196	Up	0	0	_	C	cyto	Apoplastic
Unigene36211	Up	6,57E-18	2,23E-07	_		cyto	Non-apoplastic
Unigene36268	Up	8,36E-159	1,90E-147	_		golg	Apoplastic
Unigene36307	Up	1,22E-179	3,22E-169	_		chlo	Apoplastic
Unigene36315	Up	1,53E-21	6,27E-12	_		cyto	Non-apoplastic
Unigene36335	Up	2,53E-20	9,06E-09	_		chlo	Apoplastic
Unigene36336	Up	2,90E-20	1,04E-08	C	N	mito	Non-apoplastic
Unigene36354	Up	0	0	C		chlo	Apoplastic
Unigene36379	Up	1,50E-01	1,47E+07	_		cyto_nucl	Non-apoplastic
Unigene36535	Up	0.0001461 314	0.000765	_		cyto	Apoplastic
Unigene36568	Up	2,79E-11	6,58E-01	_		cyto	Non-apoplastic

Unigene36846	Up	2,55E-02	2,40E+08	_		cyto	Non-apoplastic
Unigene36850	Up	5,22E-03	5,47E+07	_		nucl	Non-apoplastic
Unigene36901	Up	0.0001461314	0.000763	_		nucl	Apoplastic
Unigene37161	Up	0.0001461314	0.000767	_	M,N	nucl	Non-apoplastic
Unigene37172	Up	1,06E-19	4,02E-10	_		nucl	Apoplastic
Unigene37241	Up	2,99E+00	0.000186	C		chlo	Apoplastic
Unigene37294	Up	6,11E-01	4,48E+09	_		nucl	Non-apoplastic
Unigene37534	Up	7,49E-23	2,99E-11	C		nucl	Apoplastic
Unigene37828	Up	3,04E-20	1,14E-09	_	N	cyto	Non-apoplastic
Unigene37852	Up	3,71E-05	4,62E+06	_		cyto_nucl	Non-apoplastic
Unigene37856	Up	2,25E-08	3,95E+03	_	N	nucl	Non-apoplastic
Unigene37915	Up	9,30E-173	2,28E-159	_		chlo	Apoplastic
Unigene37966	Up	1,75E-53	1,64E-43	_		cyto	Non-apoplastic
Unigene37988	Up	6,35E-112	1,05E-100	C		chlo	Non-apoplastic
Unigene38087	Up	1,99E-48	1,71E-38	_		chlo	Non-apoplastic
Unigene38169	Up	1,04E+00	7,15E+09	M		cyto	Apoplastic
Unigene38264	Up	9,13E-07	1,36E+05	_		nucl	Non-apoplastic
Unigene38311	Up	8,87E-149	1,85E-136	_		nucl	Non-apoplastic
Unigene38374	Up	3,93E-10	8,55E+00	_		pero	Apoplastic
Unigene38507	Up	7,49E-23	2,98E-11	_	M,C	nucl	Non-apoplastic
Unigene38512	Up	9,75E-49	8,26E-38	_	C	chlo	Apoplastic
Unigene38530	Up	2,87E-14	8,07E-05	M		nucl	Apoplastic
Unigene38531	Up	1,32E-151	2,90E-141	_		chlo	Apoplastic
Unigene38573	Up	6,28E-12	1,45E+00	_		nucl	Non-apoplastic
Unigene38612	Up	5,04E-296	2,73E-287	_		nucl	Apoplastic
Unigene38637	Up	2,55E-02	2,38E+08	M		nucl	Non-apoplastic
Unigene38681	Up	5,21E-65	5,54E-54	_		chlo	Apoplastic
Unigene38718	Up	5,44E-21	2,09E-10	M		mito	Apoplastic
Unigene38721	Up	1,04E+00	7,13E+09	_		mito	Non-apoplastic
Unigene38765	Up	1,61E-44	1,25E-34	_	N	nucl	Apoplastic
Unigene38778	Up	1,78E-127	3,28E-117	_		chlo	Non-apoplastic
Unigene38807	Up	1,85E-28	9,77E-19	_		golg	Apoplastic
Unigene38824	Up	1,90E-25	9,01E-16	M	C,M	chlo	Non-apoplastic
Unigene38928	Up	6,60E-27	3,18E-16	C		nucl	Non-apoplastic
Unigene39100	Up	0	0	C		mito	Apoplastic
Unigene39119	Up	0	0	_		chlo	Non-apoplastic
Unigene39120	Up	2,78E-211	8,92E-199	_		extr	Apoplastic
Unigene39125	Up	4,46E-87	6,07E-76	M		nucl	Non-apoplastic
Unigene39129	Up	1,77E-243	6,81E-232	_		chlo	Apoplastic

Unigene39165	Up	6,36E-110	1,04E-98	_	N	nucl	Non-apoplastic
Unigene39169	Up	7,10E-107	1,14E-95	_		nucl	Non-apoplastic
Unigene39212	Up	7,86E-46	6,02E-34	_		nucl	Apoplastic
Unigene39228	Up	1,52E-38	1,03E-28	_		nucl	Non-apoplastic
Unigene39262	Up	1,47E-35	9,33E-26	_		cyto	Non-apoplastic
Unigene39288	Up	1,39E-29	7,55E-20	_		extr	Apoplastic
Unigene39340	Up	2,71E-08	5,12E+02	_		chlo	Apoplastic
Unigene39371	Up	2,18E-04	2,77E+06	_		cyto	Non-apoplastic
Unigene39373	Up	0.0001461 314	0.000757	_		-	Non-apoplastic
Unigene39381	Up	7,36E-02	6,43E+08	M		nucl	Non-apoplastic
Unigene39382	Up	8,52E-17	2,75E-09	_		cyto	Non-apoplastic
Unigene39386	Up	8,52E-17	2,74E-06	_		cyto	Apoplastic
Unigene39390	Up	1,74E-16	5,82E-07	M	N	nucl	Non-apoplastic
Unigene39396	Up	3,25E-02	2,99E+08	_		chlo	Apoplastic
Unigene39468	Up	6,41E-24	2,76E-13	_		nucl	Non-apoplastic
Unigene39525	Up	2,55E-02	2,40E+08	S		cyto	Non-apoplastic
Unigene39540	Up	1,50E-01	1,47E+08	C		chlo	Apoplastic
Unigene39555	Up	1,10E-05	1,80E+04	_		pero	Apoplastic
Unigene39647	Up	5,38E-06	8,24E+04	M		chlo	Apoplastic
Unigene39691	Up	5,22E-03	5,47E+05	_		nucl	Non-apoplastic
Unigene39857	Up	1,60E-07	3,10E+00	_	N	nucl	Apoplastic
Unigene39905	Up	7,36E-02	6,45E+08	_		cyto	Non-apoplastic
Unigene39933	Up	2,63E-05	3,75E+05	_		cyto	Non-apoplastic
Unigene40000	Up	2,12E-01	1,68E+08	_		chlo	Non-apoplastic
Unigene40066	Up	3,93E-10	8,51E+00	_		nucl	Non-apoplastic
Unigene40134	Up	8,61E+00	0.000476	_		nucl	Non-apoplastic
Unigene40198	Up	1,50E-01	1,47E+07	_		nucl	Non-apoplastic
Unigene40532	Up	5,60E-01	4,13E+09	C	C,M	chlo	Non-apoplastic
Unigene40600	Up	2,12E-01	1,68E+09	C	C	chlo	Non-apoplastic
Unigene40839	Up	7,36E-02	6,36E+08	_		nucl	Non-apoplastic
Unigene40952	Up	0.0001767 142	0.000903	_		cyto	Non-apoplastic
Unigene41224	Up	8,61E+00	0.000476	_		mito	Apoplastic
Unigene41262	Up	1,04E+00	7,22E+09	C		mito	Non-apoplastic
Unigene41839	Up	0.0001461 314	0.000770	_		cyto	Non-apoplastic
Unigene4423	Up	8,67E-123	1,55E-111	_		chlo	Non-apoplastic
Unigene477	Up	0	0	C		chlo	Apoplastic
Unigene5754	Up	0	0	_	C	nucl	Non-apoplastic
Unigene7586	Up	0	0	C	C,M	plas	Non-apoplastic
Unigene7732	Down	2,39E-03	2,64E+07	C	C,M,N	chlo	Non-apoplastic

Unigene7930	Up	0	0	_		chlo	Apoplastic
Unigene9158	Up	1,05E-01	1,05E+08	C		nucl	Apoplastic

E. Effector Prediction Among PstDESSPs

Table. Effector prediction, homology search and virulence prediction of PstDESSPs.

	Effector Prediction		Homology Search against Pucciniale		PHI-base Search	
ID	EffectorP 2.0	Probability	Pucciniales Homolog	e-value	PHI-Base Results	PHI-Blast e-value
CL1168.Contig1	Non-effector	0.942	PSTG_12786T0	2,00E-67		
CL1168.Contig2	Non-effector	0.964	PSTG_08840T0	1,00E-104		
CL1251.Contig1	Non-effector	0.555	PSTG_04504T0	1,00E-135	PHI:1087/F. graminearum/ unaffected pathogenicity	4.11e-19
CL1259.Contig1	Non-effector	0.887	PSTG_00428T0	3,00E-64		
CL1259.Contig3	Non-effector	0.791	PSTG_00428T1	1,00E-94		
CL1259.Contig4	Non-effector	0.649	PSTG_00428T0	1,00E-115		
CL1810.Contig1	Non-effector	0.858	PSTG_13397T1	1,00E-97		
CL1885.Contig2	Non-effector	0.787	PSTG_08590T0	2,00E-38		
CL2026.Contig2	Non-effector	0.811	PSTG_10261T0	1,00E-31		
CL2255.Contig2	Non-effector	0.611	PSTG_06874T0	1,00E-54		
CL2673.Contig1	Non-effector	0.819	PSTG_07620T0	1,00E-89		
CL2966.Contig2	Effector	0.583	PSTG_12655T1	1,00E-116		
CL3094.Contig2	Effector	0.621	PSTG_09464T0	1,00E-142	PHI:4914/Z. tritici/ Reduced virulence	4.66e-68
CL3117.Contig3	Non-effector	0.677	PSTG_07883T0	1,00E-16		
CL318.Contig1	Effector	0.911	PSTG_08979T0	1,00E-133		
CL318.Contig2	Effector	0.924	PSTG_08979T0	1,00E-127		
CL3516.Contig2	Non-effector	0.943	PSTG_11905T0	7,00E-95		
CL3737.Contig2	Non-effector	0.711	PSTG_07428T0	4,00E-25		
CL3886.Contig1	Non-effector	0.643	PSTG_04394T0	1,00E-91		
CL4066.Contig1	Unlikely effector	0.549	PSTG_04609T0	1,00E-109		
CL4121.Contig1	Non-effector	0.878	PTTG_25782T0	3,00E-19		
CL414.Contig2	Non-effector	0.613	PSTG_02885T0	1,00E-42		
CL4717.Contig1	Non-effector	0.67	PSTG_10402T0	1,00E-102		
CL4717.Contig2	Non-effector	0.67	PSTG_10402T0	1,00E-102		
CL4892.Contig2	Effector	0.865	PSTG_06134T0	4,00E-96		
CL499.Contig1	Non-effector	0.935	PSTG_15093T0	1,00E-151		
CL499.Contig2	Non-effector	0.935	PSTG_15093T0	1,00E-150		
CL5063.Contig1	Non-effector	0.575	PSTG_02340T0	2,00E-36		

CL5364.Contig2	Effector	0.81	PSTG_01880T0	1,00E-130		
CL5364.Contig3	Effector	0.793	PSTG_01881T0	1,00E-128		
CL5449.Contig1	Non-effector	0.822	PSTG_11206T0	3,00E-76		
CL5525.Contig2	Non-effector	0.582	PSTG_04068T0	1,00E-129	PHI:5236/F. oxysporum/ reduced virulence	6.29e-37
CL5553.Contig1	Effector	0.875	PSTG_10654T0	1,00E-90		
CL5553.Contig2	Effector	0.886	PSTG_10653T0	2,00E-99		
CL6430.Contig1	Non-effector	0.736	PSTG_12462T1	2,00E-63		
CL6446.Contig1	Non-effector	0.995	PSTG_06223T0	6,00E-38		
CL6446.Contig2	Non-effector	0.995	PSTG_06223T0	7,00E-38		
CL6541.Contig1	Non-effector	0.732	PSTG_16881T0	4,00E-58		
CL6640.Contig1	Non-effector	0.706	PSTG_09962T0	7,00E-59		
CL6786.Contig1	Non-effector	0.67	PSTG_02003T0	1,00E-104		
CL6816.Contig1	Effector	0.936	PSTG_01420T0	6,00E-69		
CL7100.Contig2	Effector	0.874	PSTG_05861T0	2,00E-43		
CL887.Contig1	Effector	0.829	PSTG_16009T0	2,00E-86		
CL887.Contig2	Effector	0.83	PSTG_16009T0	3,00E-92		
CL913.Contig1	Non-effector	0.957	PSTG_13745T0	5,00E-45		
Unigene10735	Non-effector	0.961	PSTG_06332T0	1,00E-109		
Unigene11912	Effector	0.708	PSTG_13268T0	9,00E-85		
Unigene12814	Non-effector	0.905	PSTG_00836T0	1,00E-106		
Unigene13023	Effector	0.933	PSTG_03450T0	7,00E-69		
Unigene16996	Non-effector	0.808	PSTG_01711T0	1,00E-165		
Unigene17495	Effector	0.833	PSTG_10917T0	1,00E-71		
Unigene19081	Non-effector	0.955	PSTG_01012T0	1,00E-136		
Unigene2221	Unlikely effector	0.539	PSTG_06810T0	1,00E-141	PHI:257/M. oryzae/ reduced virulence	1.54e-08
Unigene257	Unlikely effector	0.543	PSTG_06809T0	1,00E-144	PHI:256/M. oryzae/ reduced virulence	5.47e-06
Unigene28265	Non-effector	0.845	PSTG_02934T0	1,00E-118		
Unigene28268	Effector	0.985	PSTG_11105T0	3,00E-56		
Unigene28514	Effector	0.635	PSTG_11322T0	4,00E-83		
Unigene28527	Non-effector	0.682	PSTG_13630T0	1,00E-101		
Unigene28585	Non-effector	0.853	PSTG_04849T1	1,00E-106		
Unigene28734	Effector	0.808	PSTG_00927T0	5,00E-93		
Unigene28760	Effector	0.94	PSTG_14090T0	2,00E-73		
Unigene28808	Non-effector	0.545	PSTG_16602T0	6,00E-57		

Unigene29022	Effector	0.9	PSTG_01749T0	1,00E-139	PHI:6199/B. bassiana/ unaffected pathogenicity	6.58e-60
Unigene29087	Non-effector	0.737	PSTG_07364T0	1,00E-111		
Unigene29134	Non-effector	0.53	PSTG_17553T0	5,00E-73		
Unigene29157	Non-effector	0.58	PSTG_16911T0	1,00E-128		
Unigene29179	Effector	0.941	PSTG_04274T0	8,00E-60		
Unigene29215	Effector	0.65	PSTG_11332T0	1,00E-136		
Unigene30526	Unlikely effector	0.506	PSTG_07337T0	7,00E-57	PHI:1579/F. graminearum/ unaffected pathogenicity	2.21e-05
Unigene30571	Non-effector	0.516	PGTG_21113T0	8,00E-24		
Unigene30573	Non-effector	0.665	PSTG_06099T0	4,00E-60		
Unigene30755	Non-effector	0.796	PTTG_00486T0	3,00E-40		
Unigene30776	Non-effector	0.645	PSTG_08831T0	6,00E-71		
Unigene30809	Non-effector	0.583	PSTG_03083T0	8,00E-34		
Unigene30875	Non-effector	0.547	PSTG_03524T0	1,00E-120	PHI:2644/S. enterica/ reduced virulence	3.20e-06
Unigene30934	Non-effector	0.817	PSTG_02921T0	3,00E-91		
Unigene31019	Non-effector	0.926	PSTG_07038T0	2,00E-73		
Unigene31042	Non-effector	0.636	PSTG_01876T0	2,00E-56		
Unigene31064	Effector	0.665	PSTG_02655T0	2,00E-68		
Unigene31121	Non-effector	0.9	PSTG_13576T0	3,00E-20		
Unigene31124	Non-effector	0.642	PSTG_16218T0	6,00E-69	PHI:1566/F. graminearum/ lethal	8.72e-06
Unigene31157	Effector	0.558	PGTG_15897T0	1,00E-114		
Unigene31162	Non-effector	0.923	PSTG_14206T0	6,00E-16		
Unigene31232	Effector	0.922	PTTG_07345T0	2,00E-36		
Unigene31238	Non-effector	0.882	PGTG_11235T0	5,00E-12		
Unigene31294	Effector	0.719	PSTG_00994T0	1,00E-172		
Unigene31299	Non-effector	0.739	PSTG_01843T0	8,00E-70		
Unigene31341	Non-effector	0.634	PSTG_14884T0	1,00E-143		
Unigene31480	Non-effector	0.611	PSTG_02886T0	4,00E-11		
Unigene31503	Non-effector	0.768	PSTG_14184T0	1,00E-36		
Unigene31618	Non-effector	0.942	PSTG_17090T0	1,00E-32		
Unigene31734	Non-effector	0.638	PSTG_06452T0	1,00E-53	PHI:2533/A. fumigatus/ lethal	3.41e-28
Unigene31812	Non-effector	0.631	PSTG_13576T0	3,00E-78		
Unigene31899	Non-effector	0.9	PSTG_17264T0	6,00E-63		
Unigene31902	Effector	0.757	PSTG_07799T0	1,00E-31		

Unigene31932	Non-effector	0.852	PSTG_08524T0	2,00E-62		
Unigene31938	Effector	0.851	PSTG_04010T0	1,00E-106	PHI:383/C. albicans/ loss of pathogenicity (PHI:6412/P. striiformis/ reduced virulence)	4.27e-10 (4.14e-08)
Unigene3203	Effector	0.89	PSTG_11850T0	1,00E-114		
Unigene32092	Effector	0.889	PSTG_04048T0	5,00E-87		
Unigene32286	Effector	0.945	PSTG_01135T0	2,00E-71		
Unigene32343	Non-effector	0.922	PSTG_10617T0	5,00E-42		
Unigene32502	Non-effector	0.693	PSTG_08409T0	6,00E-23		
Unigene32923	Non-effector	0.995	PSTG_07376T0	3,00E-23		
Unigene33226	Effector	0.938	PSTG_13913T0	4,00E-24		
Unigene33242	Effector	0.57	PSTG_16262T0	2,00E-35		
Unigene33243	Non-effector	0.522	PSTG_03983T0	3,00E-64		
Unigene33278	Non-effector	0.89	PSTG_07099T0	1,00E-43		
Unigene33551	Non-effector	0.542	PSTG_05514T0	1,00E-122		
Unigene33596	Effector	0.949	PSTG_13486T0	1,00E-82		
Unigene33762	Non-effector	0.888	PSTG_10339T0	1,00E-15		
Unigene33801	Effector	0.955	PSTG_05079T0	7,00E-62		
Unigene33803	Non-effector	0.731	PSTG_13581T0	1,00E-31		
Unigene33834	Non-effector	0.785	PSTG_01770T0	1,00E-77		
Unigene33938	Effector	0.598	PSTG_00699T0	1,00E-101		
Unigene34056	Effector	0.853	PSTG_13667T0	4,00E-21		
Unigene34113	Non-effector	0.551	PSTG_01474T0	1,00E-125		
Unigene34381	Effector	0.951	PSTG_02751T0	1,00E-22		
Unigene34390	Non-effector	0.856	PSTG_11348T0	7,00E-50		
Unigene34485	Effector	0.595	PSTG_03515T0	1,00E-82		
Unigene34520	Effector	0.614	PSTG_02563T0	1,00E-88	PHI:548__PHI:2305/B. cinerea/ reduced virulence, increased virulence (hypervirulence), unaffected pathogenicity	3.38e-18
Unigene34564	Effector	0.869	PSTG_02599T0	1,00E-111		
Unigene34612	Effector	0.778	PSTG_11132T0	1,00E-125	PHI:389__PHI:1116/U. maydis/ reduced virulence, loss of pathogenicity	3.73e-41
Unigene34643	Non-effector	0.699	PSTG_01892T0	2,00E-36		
Unigene34687	Non-effector	0.577	PSTG_09177T0	1,00E-73		
Unigene34775	Effector	0.68	PSTG_15691T0	1,00E-137		
Unigene34787	Non-effector	0.57	PSTG_14510T0	1,00E-69		

Unigene35064	Non-effector	0.792	PSTG_15466T0	1,00E-113	PHI:2520/A. fumigatus/ lethal and PHI:362/ A. fumigatus/loss of pathogenicity	1.63e-19
Unigene35089	Effector	0.58	PSTG_06253T0	1,00E-110		
Unigene35403	Effector	0.872	PSTG_17053T0	1,00E-40		
Unigene35404	Effector	0.888	PSTG_17053T0	3,00E-57		
Unigene35538	Effector	0.85	PSTG_16871T0	1,00E-89	PHI:4989/B. oryzae/ reduced virulence	1.30e-30
Unigene35601	Non-effector	0.906	PSTG_02688T0	4,00E-32		
Unigene35724	Unlikely effector	0.52	PSTG_03362T0	1,00E-150		
Unigene35735	Non-effector	0.651	PSTG_01897T0	1,00E-91		
Unigene35965	Effector	0.715	PSTG_09603T0	2,00E-23		
Unigene35969	Effector	0.638	PSTG_06679T2	4,00E-88		
Unigene36059	Effector	0.8	PSTG_03254T0	2,00E-33		
Unigene36089	Non-effector	0.814	PSTG_07659T0	1,00E-41		
Unigene36187	Effector	0.671	PSTG_00676T0	5,00E-47		
Unigene36196	Non-effector	0.924	PSTG_14086T0	1,00E-168		
Unigene36211	Effector	0.554	PSTG_11460T0	1,00E-72		
Unigene36268	Non-effector	0.919	PSTG_16164T0	1,00E-145		
Unigene36307	Effector	0.843	PSTG_06305T0	1,00E-110		
Unigene36315	Effector	0.733	PSTG_16193T0	1,00E-113		
Unigene36335	Non-effector	0.704	PSTG_06435T0	1,00E-156		
Unigene36336	Non-effector	0.622	PSTG_06808T0	1,00E-123		
Unigene36354	Effector	0.77	PSTG_08969T0	1,00E-125		
Unigene36379	Non-effector	0.961	PSTG_15910T0	6,00E-99		
Unigene36535	Non-effector	0.764	PSTG_02052T0	5,00E-34		
Unigene36568	Effector	0.907	PSTG_03661T0	2,00E-94		
Unigene36846	Effector	0.746	PSTG_16782T0	1,00E-83		
Unigene36850	Non-effector	0.934	PSTG_01696T0	2,00E-55		
Unigene36901	Effector	0.892	PSTG_03879T0	4,00E-30		
Unigene37161	Non-effector	0.849	PSTG_00209T0	3,00E-75		
Unigene37172	Non-effector	0.71	PSTG_05297T0	1,00E-168		
Unigene37241	Non-effector	0.924	PSTG_11923T0	4,00E-46		
Unigene37294	Non-effector	0.852	PSTG_00238T0	1,00E-66		
Unigene37534	Non-effector	0.898	PSTG_02897T0	8,00E-34		
Unigene37828	Effector	0.692	PSTG_01576T0	3,00E-93		

Unigene37852	Non-effector	0.696	PSTG_00777T0	8,00E-26		
Unigene37856	Non-effector	0.677	PSTG_16854T0	3,00E-95	PHI:402__PHI:2224__PHI:6068/U. maydis/ loss of pathogenicity, unaffected pathogenicity, reduced virulence and PHI:193/ U. maydis/ Loss of pathogenicity	2.29e-27
Unigene37915	Effector	0.943	PSTG_09282T0	3,00E-91		
Unigene37966	Non-effector	0.5	PSTG_13342T0	2,00E-94	PHI:184/C. albicans/ reduced virulence	3.14e-06
Unigene37988	Non-effector	0.887	PSTG_16474T0	4,00E-83		
Unigene38087	Non-effector	0.902	PSTG_09694T0	3,00E-75		
Unigene38169	Non-effector	0.711	PSTG_05869T0	6,00E-76		
Unigene38264	Effector	0.86	PSTG_00708T0	1,00E-146	PHI:2920/C. graminicola/ reduced virulence	4.82e-23
Unigene38311	Effector	0.599	PSTG_04571T0	1,00E-143		
Unigene38374	Effector	0.717	PSTG_03077T0	1,00E-147	PHI:3903/A. fumigatus/ reduced virulence and PHI:2504/A. fumigatus/ unaffected pathogenicity	1.09e-54
Unigene38507	Non-effector	0.64	PSTG_09857T0	1,00E-154	PHI:6717/T. gondii/ reduced virulence	1.10e-12
Unigene38512	Effector	0.874	PSTG_14565T0	1,00E-94		
Unigene38530	Effector	0.768	PSTG_16598T0	1,00E-109		
Unigene38531	Effector	0.85	PSTG_19591T0	2,00E-33		
Unigene38573	Non-effector	0.55	PSTG_06978T0	1,00E-127	PHI:811/M. oryzae/ reduced virulence	9.02e-39
Unigene38612	Non-effector	0.683	PSTG_14728T0	1,00E-86		
Unigene38637	Effector	0.575	PSTG_00026T0	2,00E-88	PHI:2522/A. fumigatus/ lethal	2.47e-10
Unigene38681	Non-effector	0.668	PSTG_09146T0	1,00E-147		
Unigene38718	Effector	0.826	PSTG_01423T0	1,00E-141		
Unigene38721	Non-effector	0.556	PSTG_04258T0	5,00E-47		
Unigene38765	Non-effector	0.75	PSTG_00149T0	1,00E-140		
Unigene38778	Effector	0.879	PSTG_13886T0	1,00E-124		
Unigene38807	Effector	0.851	PSTG_14378T0	1,00E-128		
Unigene38824	Non-effector	0.512	PSTG_11512T0	6,00E-61		
Unigene38928	Effector	0.576	PSTG_03259T0	1,00E-130	PHI:901/U. maydis/ Unaffected pathogenicity	3.39e-69
Unigene39100	Effector	0.653	PSTG_14695T0	5,00E-85		
Unigene39119	Non-effector	0.764	PSTG_14557T0	1,00E-162		
Unigene39120	Non-effector	0.678	PSTG_17584T0	1,00E-61		

Unigene39125	Unlikely effector	0.532	PSTG_04309T0	9,00E-83	PHI:697/L. maculans/ unaffected pathogenicity	3.32e-29
Unigene39129	Effector	0.778	PSTG_12571T0	1,00E-40		
Unigene39165	Non-effector	0.626	PSTG_02113T0	3,00E-60		
Unigene39169	Non-effector	0.504	PSTG_10795T0	1,00E-148		
Unigene39212	Effector	0.775	PSTG_13167T0	1,00E-140		
Unigene39228	Effector	0.847	PSTG_00629T0	3,00E-18		
Unigene39262	Effector	0.742	PSTG_03370T0	2,00E-42		
Unigene39288	Effector	0.881	PSTG_14933T0	5,00E-53		
Unigene39340	Effector	0.859	PSTG_09382T0	4,00E-49		
Unigene39371	Unlikely effector	0.501	PSTG_06808T0	4,00E-44		
Unigene39373	Non-effector	0.658	PSTG_01847T0	8,00E-18		
Unigene39381	Non-effector	0.503	PSTG_17126T0	4,00E-44		
Unigene39382	Effector	0.564	PSTG_11331T0	5,00E-72		
Unigene39386	Effector	0.797	PSTG_16301T0	2,00E-47		
Unigene39390	Non-effector	0.892	PSTG_16882T0	1,00E-63		
Unigene39396	Effector	0.833	PSTG_08709T0	3,00E-40		
Unigene39468	Non-effector	0.686	PSTG_00445T0	7,00E-27		
Unigene39525	Non-effector	0.976	PSTG_09824T0	3,00E-72		
Unigene39540	Effector	0.704	PSTG_17409T0	1,00E-33		
Unigene39555	Effector	0.661	PSTG_07423T0	2,00E-45		
Unigene39647	Effector	0.834	PSTG_15461T0	5,00E-65		
Unigene39691	Non-effector	0.748	PSTG_00906T0	3,00E-51		
Unigene39857	Non-effector	0.943	PSTG_08566T0	1,00E-111		
Unigene39905	Effector	0.97	PSTG_15857T0	1,00E-55		
Unigene39933	Non-effector	0.738	PSTG_03634T0	3,00E-77		
Unigene40000	Non-effector	0.593	PSTG_05654T0	6,00E-43		
Unigene40066	Effector	0.86	PSTG_08556T0	1,00E-155		
Unigene40134	Non-effector	0.737	PSTG_10610T0	2,00E-28		
Unigene40198	Non-effector	0.712	PSTG_14414T0	5,00E-62		
Unigene40532	Non-effector	0.711	PSTG_13861T0	1,00E-70		
Unigene40600	Non-effector	0.813	PSTG_04328T0	2,00E-59		
Unigene40839	Effector	0.695	PSTG_09471T0	4,00E-85		
Unigene40952	Non-effector	0.662	PSTG_00107T0	6,00E-55		
Unigene41224	Effector	0.894	PTTG_27311T0	2,00E-24		

Unigene41262	Effector	0.758	PSTG_04871T0	2,00E-60		
Unigene41839	Effector	0.579	PSTG_09266T0	3,00E-54		
Unigene4423	Non-effector	0.647	PSTG_06021T0	1,00E-104		
Unigene477	Non-effector	0.917	PSTG_15335T0	4,00E-59		
Unigene5754	Effector	0.739	PSTG_08755T0	1,00E-145		
Unigene7586	Effector	0.817	PSTG_14207T0	1,00E-128		
Unigene7732	Non-effector	0.82	PTTG_26465T0	7,00E-11		
Unigene7930	Effector	0.694	PSTG_02173T0	3,00E-65		
Unigene9158	Non-effector	0.9	PSTG_01739T0	1,00E-63		

F. Comparisons of PstDESSPs with Other Reports

Table. Comparisons of PstDESSPs with other published works.

	Comparison with Cantu et al., 2013				Comparison with Xia *et al.* 2017	Comparison with Garnica et al., 2013		Comparison with Kim et al., 2016	
ID	Homologs in tribes	E-value	ISEPs	HSEPs	Avr Candidates	e-value	HSPs	e-value	SSSPs
CL1168.Contig1	PST43_03211	5e-064	Y	Y					
CL1168.Contig2	PST43_03211	9e-093	Y	Y					
CL1251.Contig1	PST21_03770	5e-066	Y	N		3e-16	Pstv_9845		
CL1259.Contig1	PST21_14352	3e-063	Y	Y		2e-66	Pstn_932		
CL1259.Contig3	PST21_14352	1e-095	Y	Y		6e-64	Pstn_932		
CL1259.Contig4	PST21_14352	4e-098	Y	Y		7e-64	Pstn_932		
CL1810.Contig1	PST21_16105	6e-070	Y	Y		2e-17	Pstv_12818	4e-19	PST130_1892
CL1885.Contig2	PST130_04598	9e-040	Y	N				3e-24	PGTG_11761
CL2026.Contig2			-	-					
CL2255.Contig2	PST877_09505	1e-048	Y	Y					
CL2673.Contig1	PST0821_12160	1e-089	Y	Y		2e-92	Pstn_13443-1		
CL2966.Contig2	PGTG_157820	2e-043	-	-					
CL3094.Contig2			-	-	YR6				
CL3117.Contig3	PST43_12993	2e-012	-	-				1e-12	PST130_21202
CL318.Contig1	PST21_16594, PST130_10538	e-118	Y	Y					
CL318.Contig2	PST21_16594	6e-079	Y	Y					
CL3516.Contig2	PST0821_10550	2e-098	Y	Y		4e-38	Pstv_15884		
CL3737.Contig2	PST130_16171	2e-034	Y	N				2e-35	PST130_12580
CL3886.Contig1	PST21_18536	9e-094	Y	Y		4e-12	Pstv_13010-1		
CL4066.Contig1	PST43_08077	e-111	Y	Y		1e-112	Pstv_15892		
CL4121.Contig1			-	-					
CL414.Contig2	PST0821_03200	2e-040	Y	Y					
CL4717.Contig1	PST43_18806	e-104	Y	Y		1e-105	Pstn_15540		
CL4717.Contig2	PST43_18806	e-104	Y	Y		1e-105	Pstn_15540		
CL4892.Contig2	PST21_14320	5e-098	Y	Y					
CL499.Contig1	PST43_19664	e-152	Y	Y		1e-153	Pstn_3128-1		
CL499.Contig2	PST43_19664	e-152	Y	Y		1e-153	Pstn_3128-1		
CL5063.Contig1									
CL5364.Contig2	PST130_05998	e-110	Y	Y		1e-119	Pstn_12353		
CL5364.Contig3	PST130_05998	e-120	Y	Y		1e-131	Pstn_12353		
CL5449.Contig1	PST0821_17588	2e-040	Y	Y		4e-14	Pstv_12818	1e-40	PST130_1892

CL5525.Contig2	PST43_00587	e-129	Y	Y					
CL5553.Contig1	PST43_20465	3e-071	Y	Y					
CL5553.Contig2	PST43_20465	e-100	Y	Y		3e-18	Pstv_15884		
CL6430.Contig1									
CL6446.Contig1	PST0821_11618	2e-040	Y	Y		3e-12	Pstn_932	8e-17	PGTG_07004
CL6446.Contig2	PST0821_11618	2e-040	Y	Y		3e-12	Pstn_932	8e-17	PGTG_07004
CL6541.Contig1	PST130_16051	2e-052	Y	Y		5e-61	Pstv_12661		
CL6640.Contig1									
CL6786.Contig1	PST21_15030	e-105	Y	Y	YR6				
CL6816.Contig1	PGTG_061710	2e-025	-	-					
CL7100.Contig2								1e-46	PST130_10324
CL887.Contig1	PST0821_09831	4e-088	Y	Y		2e-90	Pstv_8953		
CL887.Contig2	PST0821_09831	6e-094	Y	Y		6e-95	Pstv_9051		
CL913.Contig1						1e-47	Pstn_8921		
Unigene10735	PST43_19644	4e-086	Y	Y		1e-110	Pstv_8962-2		
Unigene11912	PGTG_121530	2e-037	-	-		8e-76	Pstv_671	1e-34	PGTG_12153
Unigene12814	PST21_14088	e-109	Y	Y		1e-110	Pstn_3223		
Unigene13023	PST43_10751	8e-071	Y	Y		4e-22	Pstv_9051		
Unigene16996	PST21_02044, PST43_08792	e-166	Y	Y		1e-168	Pstv_3492		
Unigene17495	PST21_17580	2e-056	Y	Y		1e-74	Pstn_16058		
Unigene19081	PST21_17885	e-114	Y	Y					
Unigene2221	PST43_01610	0.0	Y	Y		1e-21	Pstv_15614		
Unigene257	PST43_17701	e-146	Y	Y		1e-147	Pstv_15614		
Unigene28265	PST43_12937	e-120	Y	Y					
Unigene28268	PST43_03789	7e-059	-	-				1e-58	PST130_7416
Unigene28514	PST0821_05302	1e-084	Y	Y		1e-85	Pstv_3046	1e-85	PST130_9931
Unigene28527									
Unigene28585	PST43_08462	1e-108	Y	Y		1e-109	Pstn_5193		
Unigene28734	PST43_12462	1e-094	Y	Y					
Unigene28760	PST130_01890	4e-070	Y	Y		7e-14	Pstn_9466		
Unigene28808									
Unigene29022	PST21_18564	2e-016	Y	Y					
Unigene29087	PST43_10600	4e-058	Y	Y		1e-114	Pstv_4612-1		
Unigene29134	PST877_14119	4e-064	Y	Y					
Unigene29157	PST21_15640	3e-082	Y	Y		1e-130	Pstv_13398		
Unigene29179	PST130_11562	1e-058	Y	Y		1e-60	Pstv_5118		
Unigene29215	PGTG_084690	5e-073	-	-		1e-111	Pstn_15579		
Unigene30526	PST0821_01511	4e-054	Y	Y					
Unigene30571									

Unigene30573									
Unigene30755									
Unigene30776	PST0821_09560, PST21_16433	3e-072	Y	Y				3e-19	PGTG_16225
Unigene30809	PST21_19880	3e-035	Y	Y	YRTR1				
Unigene30875	PST877_16397	1e-116	Y	N		4e-13	Pstv_12298		
Unigene30934	PGTG_099670	3e-040	-	-					
Unigene31019									
Unigene31042									
Unigene31064	PST130_05630	1e-070	Y	Y					
Unigene31121						2e-31	Pstv_4624-1		
Unigene31124									
Unigene31157	PST877_16601	e-137	Y	Y		1e-57	Pstn_10436		
Unigene31162	PST0821_07134, PST130_04376	2e-017	Y	Y	YR9	2e-18	Pstn_9466		
Unigene31232									
Unigene31238									
Unigene31294	PST21_19826	e-150	Y	Y		1e-126	Pstv_15851-1		
Unigene31299								1e-72	PST130_10148
Unigene31341	PST43_17026	e-138	Y	Y		1e-147	Pstv_3225		
Unigene31480									
Unigene31503	PST21_11721	5e-038	Y	Y					
Unigene31618									
Unigene31734									
Unigene31812	PST130_08468	3e-079	Y	Y		3e-69	Pstv_4624-1	2e-80	PST130_7837
Unigene31899									
Unigene31902									
Unigene31932	PST130_02993	9e-064	Y	Y	YR9				
Unigene31938	PGTG_056670	5e-072	-	-					
Unigene3203	PST21_01496	6e-090	Y	N		2e-56	Pstn_12700		
Unigene32092	PST0821_04027, PST43_00314	1e-088	Y	Y		1e-89	Pstn_3034		
Unigene32286	PST21_19532	6e-056	Y	N		7e-36	Pstv_4021-1	5e-57	PST130_28258
Unigene32343						1e-36	Pstn_15631-1		
Unigene32502								4e-26	PST130_13247
Unigene32923	PGTG_037080	5e-021	-	-					
Unigene33226	PST130_17275	4e-057	Y	N				3e-58	PST130_9047
Unigene33242	PST43_13982	1e-025	Y	N				3e-11	PGTG_01323
Unigene33243									
Unigene33278									

Unigene33551	PST21_18096	e-123	Y	Y					
Unigene33596	PST0821_14372	3e-031	N	Y		2e-44	Pstn_5465-1		
Unigene33762									
Unigene33801	PST21_13994	6e-064	Y	Y		6e-65	Pstn_5046	4e-22	PST130_28258
Unigene33803									
Unigene33834	Mellp1_67301 fgenesh1_pg.C_scaffold_52000015	9e-018	-	-		7e-16	Pstv_3174		
Unigene33938	PGTG_180220	3e-057	-	-		1e-104	Pstv_12657		
Unigene34056	PST21_05400	6e-026	Y	Y					
Unigene34113	PGTG_016260	e-100	-	-					
Unigene34381									
Unigene34390									
Unigene34485									
Unigene34520	PST21_16249	7e-090	Y	Y		7e-91	Pstn_5784		
Unigene34564									
Unigene34612									
Unigene34643	PST21_11374	8e-038	Y	Y					
Unigene34687									
Unigene34775	PST21_15745	e-136	Y	Y					
Unigene34787									
Unigene35064									
Unigene35089	PGTG_070000	1e-086	-	-					
Unigene35403	PST21_17131	2e-011	-	-		4e-12	Pstv_16126	2e-12	PST130_10911
Unigene35404	PST21_17131	7e-013	-	-				6e-14	PST130_10911
Unigene35538	PST43_12827	6e-091	Y	Y					
Unigene35601									
Unigene35724	PST43_12562	e-137	Y	Y					
Unigene35735									
Unigene35965	PST21_12113	6e-025	Y	Y					
Unigene35969									
Unigene36059	PST21_20551	7e-032	Y	Y				5e-33	PST130_7049
Unigene36089	PST43_06959, PST21_11414	5e-043	Y	Y					
Unigene36187	PST21_05267, PST130_09542, PST43_13507	1e-031	Y	Y					
Unigene36196	PGTG_132340	e-125				1e-171	Pstn_3032		
Unigene36211	PST130_14831	8e-076	Y	Y		3e-75	Pstn_15882-1		
Unigene36268	PST21_16326	e-147	Y	Y		1e-147	Pstn_4657-1		
Unigene36307	PGTG_223100	9e-072	-	-		1e-34	Pstv_9199		

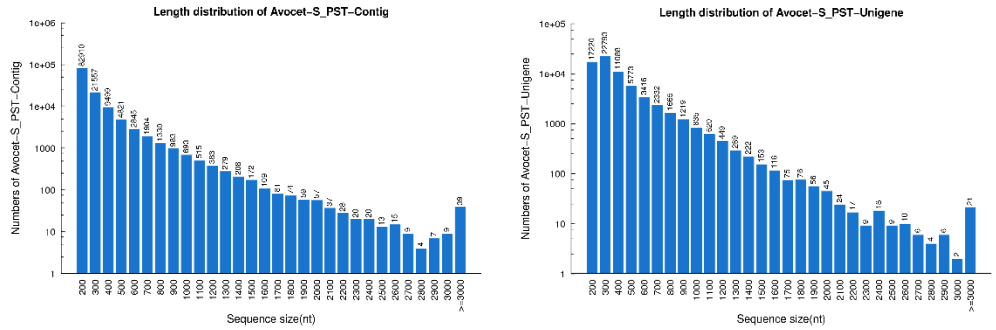
Unigene36315	Mellp1_57165 fgenesh1_kg.C _ scaffold_42000 029	6e-075	-	-					
Unigene36335	Mellp1_10487 3 fgenesh2_ pg.11_107	6e-097	-	-					
Unigene36336									
Unigene36354	PST43_08765	e-127	Y	Y		1e-12	Pstv_9199		
Unigene36379	PGTG_209280	4e-026	-	-					
Unigene36535									
Unigene36568	PST21_13514	6e-096	Y	Y		2e-97	Pstn_12695		
Unigene36846									
Unigene36850	PST130_17361 , PST43_09470	8e-057	Y	Y					
Unigene36901	PST21_13335	1e-040	Y	Y		8e-42	Pstv_16080 -1		
Unigene37161									
Unigene37172	PGTG_091200	e-106	-	-		1e-171	Pstv_12404		
Unigene37241					YR17				
Unigene37294									
Unigene37534	PST21_11684	9e-035	Y	Y				6e-36	PST130_26502
Unigene37828	PST43_04535, PST130_10194	8e-095	Y	Y		7e-96	Pstv_12696		
Unigene37852									
Unigene37856					YR9				
Unigene37915	PST21_11623	e-104	Y	Y					
Unigene37966	PST21_13429	3e-091	Y	N		4e-17	Pstv_3615- 3		
Unigene37988	PGTG_107900	2e-013	-	-					
Unigene38087	PST43_01959	5e-077	Y	Y		4e-78	Pstv_592	1e-77	PST130_1447
Unigene38169									
Unigene38264	PST130_06985	e-147	Y	N	YR6	4e-18	Pstv_12285		
Unigene38311	PST43_09436	8e-035	Y	Y		1e-146	Pstv_15883 -4		
Unigene38374	PST0821_0245 4	6e-067	Y	N					
Unigene38507	PST21_09887	2e-011	Y	Y					
Unigene38512	Mellp1_71126 estExt_fgenesh 1_ kg.C_70040	1e-020	-	-		3e-96	Pstv_4853		
Unigene38530	PGTG_156230	e-108	-	-					
Unigene38531	PST43_02992	1e-016	-	-		2e-48	Pstv_16344	1e-17	PST130_958
Unigene38573	PST130_13881	e-128	Y	Y					
Unigene38612	PST43_05098	1e-088	Y	Y		9e-47	Pstn_3128- 1		
Unigene38637									
Unigene38681	PST877_11811	4e-054	Y	N					
Unigene38718	PST43_11005	e-110	Y	Y					

Unigene38721						5e-45	Pstn_4285		
Unigene38765	PST130_05409	e-143	Y	N					
Unigene38778	PST21_08261	e-111	Y	Y					
Unigene38807	PST43_02903	e-113	Y	Y		3e-15	Pstv_9051		
Unigene38824									
Unigene38928	PGTG_027770	e-112							
Unigene39100	PST0821_0864 8	1e-077	Y	Y					
Unigene39119	PGTG_086380	5e-053	-	-		1e-165	Pstn_3099		
Unigene39120	PST21_04062	1e-062	Y	Y					
Unigene39125									
Unigene39129	PST0821_0416 6	5e-028	-	-					
Unigene39165	PST0821_0479 8	7e-059	Y	N				6e-60	PST130_26236
Unigene39169	PST43_03505	e-150	Y	Y		1e-12	Pstv_12298		
Unigene39212	PST43_17014	e-142	Y	Y					
Unigene39228	PST0821_0649 6	7e-077	Y	Y		7e-78	Pstv_15878		
Unigene39262	PST21_14100	3e-017	Y	Y		2e-17	Pstv_5924	2e-83	PST130_24658
Unigene39288	PST21_18506	3e-055	Y	N		1e-24	Pstn_5193		
Unigene39340	PST130_13044	1e-050	Y	Y		1e-51	Pstv_4607		
Unigene39371									
Unigene39373	PST0821_0387 2, PST43_18621	3e-019	Y	Y					
Unigene39381									
Unigene39382						2e-76	Pstv_3060		
Unigene39386						2e-50	Pstn_409		
Unigene39390	PST130_16050	1e-056	Y	Y		2e-66	Pstv_4804-1		
Unigene39396	PST21_15706	9e-042	Y	N					
Unigene39468									
Unigene39525									
Unigene39540									
Unigene39555									
Unigene39647	PST877_00513	3e-056	Y	N		5e-14	Pstn_3451		
Unigene39691	PST877_11879 , PST21_20122	1e-052	Y	Y					
Unigene39857	PST21_19352	e-103	Y	Y		1e-114	Pstv_4622-2		
Unigene39905	PST43_05168	2e-057	Y	Y		2e-58	Pstv_3161-1	1e-58	PST130_15220
Unigene39933	PST43_06790	5e-079	Y	Y		5e-80	Pstv_3710	8e-80	PST130_28870
Unigene40000	PST43_02946	2e-044	-	-					
Unigene40066	PST21_11390	e-117	Y	Y		1e-133	Pstv_15897		
Unigene40134	PST877_11462 ,	8e-030	Y	Y					

	PST0821_05552								
Unigene40198	PST43_11457, PST130_00860, PST21_12116	3e-058	Y	Y		5e-40	Pstv_12292		
Unigene40532	PST877_03536	2e-071	Y	Y					
Unigene40600	PST130_15913	5e-058	Y	Y					
Unigene40839									
Unigene40952									
Unigene41224									
Unigene41262	PST0821_05277	2e-061	Y	Y		2e-62	Pstv_15885-1		
Unigene41839	PST0821_03901	8e-057	Y	Y				6e-15	PGTG_17995
Unigene4423	PST43_00805	e-106	Y	Y		1e-107	Pstv_15884		
Unigene477	PST877_17725	4e-064	Y	Y		1e-65	Pstv_12818	9e-12	PGTG_12844
Unigene5754	PST43_10036	e-147	Y	Y					
Unigene7586	PST130_04375	e-130	Y	Y	YR9	3e-87	Pstv_4516-2		
Unigene7732									
Unigene7930	PST43_01930	3e-071	Y	Y		3e-72	Pstv_5004-1	2e-72	PST130_3448
Unigene9158	PST43_04767, PST877_00240	2e-065	Y	N		7e-62	Pstn_16224-1		

G. Length Distribution of Contigs and Unigenes

A **B**



C **D**

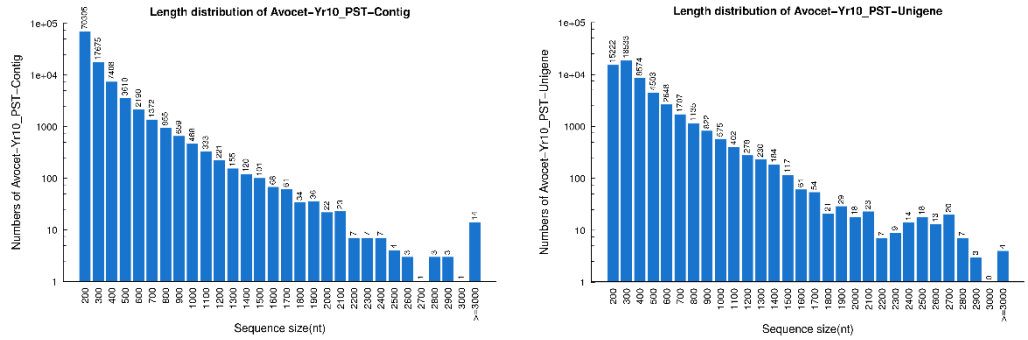


Figure: Length distribution of contigs and unigenes after assembly process. Graphical representation of length distribution of **A)** Contigs of AvocetS_Pst, **B)** Unigenes of AvocetS_Pst, **C)** Contigs of AvocetYR10_Pst and **D)** Unigenes of AvocetYR10_Pst.

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name	: Özketen, Ahmet Çağlar
Nationality	: Turkish (TC)
Date and Place of Birth	: 13.09.1988, İzmir
Phone	: +90 554 325 86 17
E-mail	: caglar.ozketen@metu.edu.tr

EDUCATION

Degree	Institution	Year of Graduation
MS	METU Biotechnology	2013
BS	METU Molecular Biology and Genetics	2011
High School	MEV İzmir Private Science High school, İzmir	2006

WORK EXPERIENCE

Year	Place	Enrollment
2011-Present	METU Akkaya Resarch Group	Researcher

FOREIGN LANGUAGES

Advanced English

PUBLICATIONS

• Dagvadorj B, **Ozketen AC**, Andac A, Duggan C, Bozkurt TO, Akkaya MS. 2017. A Puccinia striiformis f. sp. tritici secreted protein activates plant immunity at the cell surface. Scientific Reports 7, 1141.