USE OF THESES

This copy is supplied for purposes of private study and research only. Passages from the thesis may not be copied or closely paraphrased without the written consent of the author.
Strategies for wheat stripe rust pathogenicity identified by “omics” technologies

Diana Patricia Garnica Moreno

A thesis submitted for
The degree of Doctor of Philosophy
of
The Australian National University

July 2014
ACKNOWLEDGEMENTS

I would like to immensely thank my supervisor Dr John Rathjen, for opening the doors of his lab to me and entrusting me such an interesting research project. I deeply thank him for all his guidance and support every day of my PhD, for always having time to discuss ideas and for encouraging me to keep going. Thanks to my co-supervisor Dr Peter Dodds, for his valuable input in the development of my project and for the time dedicated to have open science discussions. Also, thanks to Dr Narayana Upadhyaya for all the support along my project, for always having time for me to solve little or big questions and for believing in me as scientist.

Thanks to Dr Colin Wellings and Dr Robert Park for providing me the biological material and the basic knowledge to work with wheat rust pathogens. A special thanks to Dr Harpreet Vorah and Mr Michael Devoy for all their time, collaboration and support with the flow cytometry and microscopy experiments. Thanks to Dr Carsten Kulheim for all your support with the analysis of transcriptomic data. Thanks you to the CSIRO program Y technical personnel, Kim, Xiaodi, Sutha and Robin and also Dr Rohit Mago for all your support and expertise.

Many thanks to my lab mates Will, Brendon, Veronica and Isabel, for the valuable scientific discussions we had, for your intellectual input to my project and specially for their friendship, you really made my time in the lab very enjoyable. To my friends Clarissa and Yvonne, thank you very much for your daily support, for your sincere friendship, nothing would have been the same without you.

Thank you very much to my parents, my sister and my grandmother that despite the distance, always were there to cheer me up and for encouraging me to move forward.

Last but not least, thanks to my rock Mario and my inspiration Isabella, I do not have enough words to express my gratitude for your support.

This thesis was made possible through support from an ANU HDR Merit Scholarship.
Striped rust is a major constraint to wheat production worldwide. The causal agent is the fungus *Puccinia striiformis* f.sp. *tritici* (*Pst*), a close relative of the pathogens which cause the devastating wheat stem and leaf rust diseases. During infection, the fungus creates a specialized cellular structure within host cells called the haustorium which invades the cells of the wheat leaf, where it can extract nutrients such as sugars directly from the host. In addition, the haustorium is thought to secrete proteinaceous virulence molecules called ‘effectors’, which are suspected to manipulate the physiological and immune responses of host cells during infection. Despite this broad outline, the molecular events that underlie host colonization and the effectors proteins produced by *Pst* are largely unknown. In my PhD, I extensively investigated *Pst* using transcriptomics and proteomics techniques to obtain a better understanding of how the pathogen establishes a compatible interaction with its host, and to identify the effector proteins that are synthesised and secreted during infection. First, using next generation sequencing techniques (454 and Illumina), the transcriptomes of two contrasting pathogenic stages (*in vitro* germinated spores and haustoria from the infectious stage) were generated, assembled *de novo* and annotated extensively. A digital gene expression analysis revealed many differentially expressed genes which highlight important metabolic differences between the developmental stages, and provide insight into their different roles during infection. Spores express mainly metabolic pathways to derive energy from non-carbohydrate sources, necessary for growth and development. Conversely, haustoria deploy molecular machinery to exploit the abundant nutrients available from the host, and focus on energy production and biosynthetic pathways to support fungal growth and spore production. Further analysis of the haustorial transcriptome allowed me to identify the first comprehensive set of potential effector candidate genes of *Pst*, comprising 437 genes, with two thirds of these up-regulated in haustoria by comparison to germinated spores. Using a bacterial system to synthesise and deliver proteins encoded by effector gene candidates, a subset of these genes was cloned and used to establish two strategies for their functional characterization. The first aimed to test if these proteins could be recognised by wheat resistance genes leading to the induction of defense responses, and the second tested their capacity to inhibit cell death triggered by a necrotic toxin. Although no recognized effectors were identified in this screen, two effector candidates partially inhibited plant cell death in the second scheme. In parallel, I developed a method combining density gradients and automated flow cytometry to isolate highly purified haustoria. Haustoria purified by this method were used successfully for proteomics analysis. Proteomics data sets from haustoria, germinated and ungerminated spores were generated, and a preliminary analysis was undertaken to determine the presence of candidate effector
proteins as well as non-effector proteins in each tissue. More than 3,000 proteins were validated in the proteomic data, including 150 candidate effectors. Differences between the transcriptomic and proteomic data suggested that the synthesis and deployment of some effector proteins could occur at different spatiotemporal sites, and could perhaps have destinations other than direct secretion to the host cell cytoplasm. Together, these studies have substantially increased our knowledge of Pst effectors and have provided insights into the pathogenic strategies of this important organism, while opening new avenues of research with immense potential in the design of novel disease control strategies.
### ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-DE</td>
<td>SDS-PAGE in a 2 dimension format</td>
</tr>
<tr>
<td>7-AAD</td>
<td>7-amino-actinomycin D</td>
</tr>
<tr>
<td>aa</td>
<td>amino acid</td>
</tr>
<tr>
<td>Avr</td>
<td>Avirulence gene</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Search Tool</td>
</tr>
<tr>
<td>BLASTn</td>
<td>BLAST search of nucleotide databases using a nucleotide query</td>
</tr>
<tr>
<td>BLASTx</td>
<td>BLAST search of protein databases using a translated nucleotide query</td>
</tr>
<tr>
<td>BLASTp</td>
<td>BLAST search of protein databases using a protein query</td>
</tr>
<tr>
<td>B2G</td>
<td>Software BLAST2GO</td>
</tr>
<tr>
<td>ConA</td>
<td>Concanavalin A</td>
</tr>
<tr>
<td>ConA-488</td>
<td>Concanavalin A, Alexa Fluor 488 Conjugate</td>
</tr>
<tr>
<td>dNTPs</td>
<td>Deoxynucleotide Triphosphates</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EDB</td>
<td>Effector database</td>
</tr>
<tr>
<td>EHM</td>
<td>Extrahaustorial membrane</td>
</tr>
<tr>
<td>EHx</td>
<td>Extra haustorial matrix</td>
</tr>
<tr>
<td>EST</td>
<td>Expressed sequence tags</td>
</tr>
<tr>
<td>ETI</td>
<td>Effector-triggered immunity</td>
</tr>
<tr>
<td>EV</td>
<td>Empty vector</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence-activated cell sorting</td>
</tr>
<tr>
<td>FDR</td>
<td>False discovery rate</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescent protein</td>
</tr>
<tr>
<td>GH</td>
<td>Glycoside hydrolases</td>
</tr>
<tr>
<td>GO</td>
<td>Gene ontology</td>
</tr>
<tr>
<td>HMC</td>
<td>Haustorial mother cell</td>
</tr>
<tr>
<td>HR</td>
<td>Hypersensitive response</td>
</tr>
<tr>
<td>HSP</td>
<td>Haustorial secreted proteins</td>
</tr>
<tr>
<td>IB</td>
<td>Isolation buffer</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-bertani broth</td>
</tr>
<tr>
<td>LC</td>
<td>Liquid chromatography</td>
</tr>
<tr>
<td>Mlp</td>
<td>Melampsora larici-populina</td>
</tr>
<tr>
<td>MS</td>
<td>Mass spectrometry</td>
</tr>
<tr>
<td>NGS</td>
<td>Next-generation sequencing</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td>P20</td>
<td><em>in situ</em> density gradient from solutions of 20% Percoll (v/v)</td>
</tr>
<tr>
<td>P30</td>
<td><em>in situ</em> density gradient from solutions of 30% Percoll (v/v)</td>
</tr>
<tr>
<td>PAMP</td>
<td>Pathogen-associated molecular patterns</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal component analysis</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>Pgt</td>
<td><em>Puccinia graminis</em> f.sp. <em>tritici</em></td>
</tr>
<tr>
<td>PIG</td>
<td>Planta-induced fungal genes</td>
</tr>
<tr>
<td>PPS</td>
<td>Pentose phosphate shunt</td>
</tr>
<tr>
<td>PRR</td>
<td>Pattern recognition receptors</td>
</tr>
<tr>
<td>Pst</td>
<td><em>Puccinia striiformis</em> f.sp. <em>tritici</em></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Pt</td>
<td><em>Puccinia triticina</em></td>
</tr>
<tr>
<td>PTI</td>
<td>PAMP-triggered immunity</td>
</tr>
<tr>
<td>R</td>
<td>Resistance gene</td>
</tr>
<tr>
<td>RPKM</td>
<td>Reads per kilobase of exon model per million mapped reads</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcriptase-PCR</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium-dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SP</td>
<td>Signal peptide</td>
</tr>
<tr>
<td>TCA</td>
<td>Tricarboxylic acid cycle</td>
</tr>
<tr>
<td>TCA</td>
<td>Trichloroacetic acid</td>
</tr>
<tr>
<td>TDB</td>
<td>Transcriptome database</td>
</tr>
<tr>
<td>TTSS</td>
<td>Type three secretion system</td>
</tr>
</tbody>
</table>
CONTENTS

Declaration ...................................................................................................................................... ii
Acknowledgements ......................................................................................................................... iii
Abstract ........................................................................................................................................... iv
Abbreviations .................................................................................................................................. vi
List of figures .................................................................................................................................... xii
List of tables ..................................................................................................................................... xiii

CHAPTER 1
INTRODUCTION TO WHEAT STRIPE RUST

1.1. WHEAT RUST FUNGI ............................................................................................................... 2
  1.1.1. Economic impact of the wheat rusts ................................................................. 4
1.2. STRIPE RUST: UNDERSTANDING THE PATHOGEN AND THE DISEASE IN WHEAT .......... 5
  1.2.1. Life cycle ................................................................. 5
  1.2.2. Infection of wheat by Pst uredospores ............................................................ 7
1.3. BIOTROPHIC INTERFACES .................................................................................................... 8
  1.3.1. The haustorial complex .............................................................................. 9
  1.3.2. Nutrient uptake by rust haustoria ............................................................... 12
  1.3.3. Additional roles of haustoria ................................................................. 13
1.4. BIOTROPHIC EFFECTORS .................................................................................................. 14
  1.4.1. Extracellular effectors ............................................................................. 16
  1.4.2. Cytoplasmic effectors ............................................................................. 17
  1.4.3. Effector identification and validation .................................................... 19
1.5. PLANT DEFENCE .................................................................................................................. 22
  1.5.1. Resistance against stripe rust in wheat .................................................... 24
1.6. SEQUENCING TECHNOLOGIES ...................................................................................... 25
  1.6.1. Pst genomic resources ........................................................................ 26
1.7. PROJECT AIMS AND EXPERIMENTAL APPROACH ........................................................ 28

CHAPTER 2
PREDICTION AND CHARACTERIZATION OF Pst EFFECTORS

2.1. INTRODUCTION ...................................................................................................................... 29
2.2. MATERIALS AND METHODS .............................................................................................. 31
  2.2.1. Fungal and plant material ........................................................................ 31
  2.2.2. Infection of wheat with Pst ...................................................................... 32
2.2.2.1. Strain purification by single pustule isolation ............................................ 32
2.2.2.2. Plant infection and collection of infected tissue ........................................ 33
2.2.3. Haustoria isolation and transcriptome sequencing (RNA-seq) .................... 33
  2.2.3.1. Haustoria isolation by Concanavalin A (ConA) affinity chromatography ...
    2.2.3.1.1. Coupling of CNBr-activated Sepharose 6MB to ConA .................. 34
    2.2.3.1.2. Regeneration of column after use ........................................... 34
    2.2.3.1.3. Isolation of haustoria from Pst-infected wheat leaves ............ 34
  2.2.3.2. Total RNA isolation from purified haustoria and germinated spores ...... 35
  2.2.3.3. mRNA isolation and pyrosequencing .................................................... 36
2.2.4. De novo assembly of haustoria and germinated spores transcriptomes, and prediction of effector gene candidates ................................................................. 36
2.2.5. Characterization of candidate effector genes ................................................ 37
  2.2.5.1. Validation of expression of candidate effector genes by reverse
           transcriptase-PCR (RT-PCR) ................................................................. 37
  2.2.5.2. Annotation of candidate effector genes .............................................. 37
  2.2.5.3. Cloning of candidate effector genes into a specialized vector
           (pEtHAn) for secretion via the type-III secretion system of
           Pseudomonas fluorescens ........................................................................ 38
    2.2.5.3.1. Cloning of candidate effector genes into the Gateway
               entry vector pENTR/D-TOPO ......................................................... 38
    2.2.5.3.2. Recombination of candidate effector genes into the
               destination expression vector pNR526-C2AG3A ......................... 39
    2.2.5.3.3. Transformation of Pseudomonas fluorescens Pf0-1
               (pEtHAn) bacteria by triparental mating ....................................... 39
    2.2.5.3.4. Activation of the P. fluorescens type-III secretion
               system for infiltration assays ......................................................... 40
  2.2.5.4. Effector recognition assay .................................................................. 40
  2.2.5.5. Plant cell death inhibition assay ........................................................ 41

2.3. RESULTS ........................................................................................................... 41
  2.3.1. Pst isolate purification .............................................................................. 41
  2.3.2. Isolation of haustoria .............................................................................. 43
  2.3.3. RNA isolation from haustoria and spores and purification of mRNA ...... 44
  2.3.4. Assembly of contigs from haustoria and spores RNA-seq data, and prediction
         of effector genes ....................................................................................... 46
  2.3.5. Validation of expression of HSP genes ................................................... 49
  2.3.6. Expression profiles of HSP genes ........................................................... 50
  2.3.7. Annotation of HSP genes ........................................................................ 51
    2.3.7.1. Length and cysteine content ........................................................... 51
    2.3.7.2. Functional classification .................................................................. 54
  2.3.8. Potential recognition of Pst HSPs by wheat tester lines ......................... 55
  2.3.9. Assays for inhibition of plant cell death (PCD) by Pst HSPs .................. 59

2.4. DISCUSSION ..................................................................................................... 65
  2.4.1. Purification of Pst haustoria from infected leaf tissue ............................. 65
  2.4.2. Prediction and expression analysis of the Pst effector candidate genes ....... 67
  2.4.3. Sequence analysis of the predicted Pst HSPs ......................................... 68
  2.4.4. Functional characterization of Pst HSPs ................................................. 70
CHAPTER 3
TRANSSCRIPTOME ANALYSIS OF Pst

3.1. INTRODUCTION ........................................................................................................................ 73
3.2. MATERIALS AND METHODS ..................................................................................................... 75
  3.2.1. Semi-purification of Pst haustoria using Percoll gradients .................................................. 75
  3.2.2. RNA isolation and sequencing from Pst haustoria and germinated spores ...................... 76
  3.2.3. Reference transcriptome assembly and BLAST2GO and BLASTx annotation .................. 76
  3.2.4. Digital expression analysis .................................................................................................. 77
3.3. RESULTS AND DISCUSSION ....................................................................................................... 78
  3.3.1. Assembly of a reference transcriptome for gene expression analysis ................................. 78
  3.3.2. Isolation of Pst haustoria using Percoll gradients, and RNA extraction from haustoria and germinated spores ................................................................. 79
  3.3.3. Illumina sequencing ............................................................................................................ 81
  3.3.4. Digital expression analysis ................................................................................................ 82
  3.3.5. Functional classification of the transcriptome sequences .................................................. 87
    3.3.5.1. Cell proliferation ............................................................................................................. 88
      3.3.5.1.1. DNA replication and cell cycle .................................................................................. 88
      3.3.5.1.2. Cell wall modification enzymes .............................................................................. 90
    3.3.5.2. Transporters .................................................................................................................. 92
      3.3.5.2.1. Amino acid and oligopeptide transporters .............................................................. 92
      3.3.5.2.2. ABC and MFS transporters .................................................................................... 93
      3.3.5.2.3. Sugar transporters .................................................................................................. 94
      3.3.5.2.4. Phosphate transporters .......................................................................................... 95
      3.3.5.2.5. Nicotinic acid transporters .................................................................................... 96
    3.3.5.3. Primary metabolism ....................................................................................................... 97
      3.3.5.3.1. Energy use - oxidative phosphorylation .................................................................. 97
      3.3.5.3.2. Fatty acid metabolism ............................................................................................ 97
      3.3.5.3.3. Pentose phosphate shunt (PPS) .............................................................................. 99
      3.3.5.3.4. Thiamine biosynthesis .......................................................................................... 101
      3.3.5.3.5. Nitrogen metabolism ............................................................................................ 102
      3.3.5.3.6. Sulfur metabolism ................................................................................................ 102
    3.3.5.4. Protection against host defences .................................................................................... 103
      3.3.5.4.1. Glutathione metabolism ......................................................................................... 103
      3.3.5.4.2. Cytochrome P450s ................................................................................................. 104
    3.3.5.5. Transcriptions factors .................................................................................................. 104
    3.3.5.6. Signal transduction ....................................................................................................... 105
3.4. CONCLUSION ............................................................................................................................ 107

CHAPTER 4
PURIFICATION OF Pst HAUSTORIA

4.1. INTRODUCTION ........................................................................................................................ 109
4.2. MATERIALS AND METHODS ................................................................. 113
  4.2.1. Haustoria enrichment on Percoll density gradients ......................... 113
    4.2.1.1. Processing of Pst-infected tissue prior to density gradient 
              centrifugation ................................................................. 113
    4.2.1.2. General steps for the preparation of discontinuous Percoll gradients
    4.2.1.3. General steps for the preparation of self-formed continuous 
              Percoll gradients ............................................................................. 114
  4.2.2. Purification of haustoria from enriched fractions by fluorescent-activated 
          cell sorting (FACS) ....................................................................... 114
    4.2.2.1. Concanavalin A- Alexa 488 staining ..................................... 114
    4.2.2.2. Fluorescent-Activated Cell Sorting of ConA-488 stained haustoria ...... 114
  4.2.3. Viability assays on sorted haustoria ........................................... 115
    4.2.3.1. 7-amino actinomycin D staining of dead cells ......................... 115
    4.2.3.2. CellTracker Orange CMRA viability test ................................ 116
4.3. RESULTS .......................................................................................... 116
  4.3.1. Enrichment of haustoria by discontinuous Percoll gradients .............. 116
  4.3.2. Enrichment of haustoria using self-formed continuous Percoll gradients 121
  4.3.3. Haustoria purification by flow cytometry ....................................... 125
  4.3.4. Viability tests on sorted haustoria .............................................. 127
    4.3.4.1. Haustoria viability tested by 7-amino actinomycin D (7-AAD) dye 
              exclusion .................................................................................. 127
    4.3.4.2. Haustoria viability tested using CellTracker CMRA ................... 131
4.4. DISCUSSION ...................................................................................... 137
  4.4.1. Haustoria isolation by discontinuous and continuous Percoll gradients .... 137
  4.4.2. Purification of haustoria by fluorescence-activated cell sorting (FACS) and 
          evaluation of their viability post sorting ...................................... 141

CHAPTER 5
PROTEOMICS OF Pst SPORES AND HAUSTORIA AS A TOOL TO IMPROVE THE 
INVESTIGATION OF EFFECTORS

5.1. INTRODUCTION ................................................................. 147
5.2. MATERIALS AND METHODS ............................................................. 151
  5.2.1. Sample processing and protein extraction from ungerminated and germinated 
         Pst uredospores .............................................................................. 151
  5.2.2. OFFGEL fractionation of spore proteins ...................................... 152
  5.2.3. Protein extraction from FACS-purified Pst haustoria ....................... 152
  5.2.4. Capillary liquid chromatography and mass spectrometry of Pst peptides 
         ................................................................. 152
  5.2.5. Mass spectrometry data analysis ............................................... 153
  5.2.6. Assessment of effector candidate gene expression in ungerminated spores 
         by reverse transcriptase-PCR (RT-PCR) .................................... 153
5.3. RESULTS .......................................................... 154
  5.3.1. Extraction of proteins from Pst spores and FACS-purified haustoria .... 154
  5.3.2. Pst transcriptomic resources used for protein identification .......... 156
CHAPTER 6
GENERAL DISCUSSION

6.1. PERSPECTIVES ................................................................. 166
6.2. FUTURE DIRECTIONS ................................................. 168

APPENDICES ................................................................................. 173
REFERENCES .................................................................................. 179
PUBLISHED MATERIAL ................................................................. 210
LIST OF FIGURES

Figure 1.1. Lesions caused by wheat rust pathogens
Figure 1.2. Life cycle of Puccinia striiformis
Figure 1.3. The infection process of rust pathogens on wheat
Figure 1.4. Plant defence
Figure 2.1. Single pustule isolation
Figure 2.2. Morocco wheat seedlings infected with Pst 104 E137 A-.
Figure 2.3. Stripe rust haustoria isolated by affinity chromatography
Figure 2.4. Staining of isolated Pst haustoria with ConA and WGA.
Figure 2.5. mRNA analysis with the Agilent 2100 bioanalyzer
Figure 2.6. Number of 454 reads per assembled contig for spores and haustorial samples
Figure 2.7. Length of contigs for spores and haustorial samples assembled from 454 data
Figure 2.8. Coverage of contigs assembled from 454 data from spores and haustoria samples
Figure 2.9. HSPs expression patterns studied by RT-PCR
Figure 2.10. HSPs expression patterns
Figure 2.11. Effector candidates grouped according to distribution of cysteines in the predicted protein sequence
Figure 2.12. Schematic representation for delivery of Pst HSPs via the TTSS of P. fluorescens Pf0-1 (EThAn)
Figure 2.13. AvrM-M recognition in tobacco
Figure 2.14. Standardization of optimal ToxA concentration for induction of cell death on non-infected wheat cv. Grandin.
Figure 2.15. Inhibition of ToxA-induced plant cell death by Pst infection
Figure 2.16. Inhibition of ToxA-induced cell death by bacterial delivery of Pst HSPs
Figure 3.1. Bioanalyzer electropherograms of total RNA
Figure 3.2. Average Phred quality score per base for haustoria and spores libraries before quality control
Figure 3.3. Principal Component Analysis of RNAseq replicate samples
Figure 3.4. Differential gene expression in germinated spores and haustoria of Pst
Figure 3.5. Mapping of Illumina reads from haustoria and germinated spores against transcripts from the reference set
Figure 3.6. Comparative ontological analysis of transcripts with statistically significant changes in expression between haustoria and germinated spores
Figure 3.7. Heatmaps representing relative levels of gene expression classified by metabolic categories
Figure 3.8. Metabolic processes in germinated spores
Figure 3.9. Metabolic processes in haustoria
Figure 4.1. Generalized illustration of infection structures of rust fungi
Figure 4.2. Distribution of Pst haustoria on a discontinuous Percoll gradient
Figure 4.3. Measurement of the haustoria:chloroplast ratio in a three-step discontinuous Percoll gradient
Figure 4.4. A three-step Percoll gradient for haustoria enrichment
Figure 4.5. Self-formed continuous Percoll gradient for haustoria enrichment
Figure 4.6. Distribution of Pst haustoria on self-formed continuous 30% Percoll gradients: haustoria distribution and gradient density
Figure 4.7. Purification of *Pst* haustoria by FACS, and microscopic analysis of pre- and post-sorted samples
Figure 4.8. Pre-sorted haustoria treated with 7-AAD
Figure 4.9. Pre-sorted haustoria treated with extran and then 7-AAD
Figure 4.10. Microscopic analysis of pure haustoria treated with 7-AAD
Figure 4.11. FACS analysis of pure haustoria tested for viability with 7-AAD
Figure 4.12. Standardization of CMRA concentration to assess haustorial viability in pre-sorted samples
Figure 4.13. CMRA viability test on pre-sorted haustoria samples previously subjected to heat and chemical treatments
Figure 4.14. CMRA viability test applied to haustoria-enriched samples treated previously with 70% isopropanol
Figure 4.15. CMRA viability test applied on sorted haustoria
Figure 5.1. SDS-PAGE of protein extracts from germinated and ungerminated spores
Figure 5.2. Proteins in *Pst* haustoria and spores
Figure 5.3. Effector gene expression in ungerminated *Pst* spores and haustoria

**LIST OF TABLES**

Table 1.1. Physical characteristics used to differentiate the wheat rust diseases
Table 2.1. *Pst* HSP classes with characteristic Cys patterns
Table 2.2. Pathotyping *Pst* strain 104 E137 A- to select resistant wheat cultivars
Table 3.1. Illumina sequencing data before and after mapping against the unigene reference
Table 4.1. Effect of centrifugation time on the distribution and purity of haustoria on a three step Percoll gradient
Table 4.2. Comparison of the haustoria isolation methods applied on *Pst* infected tissue
Table 5.1. Proteins identified from *Pst* proteomic data generated from germinated spores, ungerminated spores and FACS-purified haustoria
Chapter 1

Introduction to wheat stripe rust

The rust fungi are a large monophyletic group of obligate biotrophic organisms that grow, reproduce and complete their life cycle only on live plant tissue (Kolmer et al., 2009). Their name comes from the rust coloured pustules that erupt on the plant epidermis, which are full of spores produced clonally on the hosts during sporulation. There are approximately 7,000 species of rust fungi, all in the phylum Basidiomycota, order *Pucciniales*, all of which are highly specialized parasites of a broad range of plant host species (Kirk et al., 2008). Throughout history, rust fungi have caused famines and have severely affected the economies of many countries throughout the destruction of crops grown for staple human foods such as wheat, oats, barley and rye (Bushnell and Roelfs, 1985). Other plants including legumes, ornamentals and trees of economic importance such as coffee and apples, can also be attacked by rusts (Agrios, 2005). Rust fungi have evolved such that most of them have a very narrow host range of plant species on which they can infect, colonize and reproduce. There is often a great deal of variability within a single rust species in the ability to attack different genotypes within one host species (Kolmer et al., 2009). Such a level of specialization causes an evolutionary arms race between the pathogen and the host, making the rust fungi a fascinating genetic model. Rust pathogens commonly infect leaves and stems, although occasionally infection can be seen in floral parts and fruits (Agrios, 2005). The biotrophic nature of rust fungi makes them difficult to study as they can only be cultured *in vitro* with great difficulty, and the methods for genetic transformation
of these fungi are still in the early stages (Kern, 1956, Lawrence et al., 2010, Schillberg et al., 2000, Djulic et al., 2011).

1.1. WHEAT RUST FUNGI

Rust diseases of wheat are among the oldest plant diseases known to man, as references to wheat rust can be found in the Bible and the classical literature of ancient Greece and Rome (Chester, 1946). Since the discovery of rusts, studies on their life cycle and management have allowed the development of control methods to reduce the impact of these diseases (Carver, 2009). Although today epidemics are less common worldwide due mainly to the use of resistant wheat cultivars, these diseases are acknowledged to pose a continuous threat to food security (McIntosh et al., 1995). The continuing evolution of virulent rust races and the persistence of rust as a significant disease in wheat can be attributed to specific characteristics of these pathogens. The capacity to produce a large number of spores which can be easily disseminated over long distances by wind, and the ability to change genetically in response to the deployment of rust resistant wheat cultivars, facilitates the emergence of new races that are more aggressive on resistant wheat lines (Marsalis and Goldberg, 2006).

There are three rust diseases that occur on wheat: stem, leaf and stripe rust (Figure 1.1). These diseases are each caused by a particular species of the fungal genus *Puccinia*, which belong to Basidiomycetes, order Pucciniales. The genus *Puccinia* currently contains approximately 4000 species (Fellers et al., 2005). Stem rust (*Puccinia graminis* f.sp. *tritici*, *Pgt*), leaf rust (*Puccinia triticina*, *Pt*) and stripe rust (*Puccinia striiformis* f.sp. *tritici*, *Pst*), produce similar disease symptoms on host plants and have similar requirements for infection (Marsalis and Goldberg, 2006). Infection can occur on any aerial plant organ, ultimately leading to the production of pustules that contain thousands of dry yellow-orange to reddish-brown uredospores (Wolf et al., 2010). These pustules give infected plants a “rusty” look. Characteristics such as which plant parts are affected, or arrangement of the pustules on plant surfaces, allow visual discrimination of these diseases (Wolf et al., 2010). Other characteristics including lesion size, shape, and colour, can help to confirm the diagnosis or separate the more similar diseases, although sometimes they are less clear. Figure 1.1 and Table 1.1 establish a comparison between the three wheat rusts based on their location on the host, and the lesions they cause.
Figure 1.1. Lesions caused by wheat rust pathogens. A-B Stem rust, C-D Leaf rust, E-F. Stripe rust. Images taken from (Wolf et al., 2010)

<table>
<thead>
<tr>
<th>Synonym</th>
<th>Stem rust</th>
<th>Leaf rust</th>
<th>Stripe rust</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parts of plant infected</td>
<td>Primarily on stems but can also be found on leaves, sheaths, glumes, awns and even seeds.</td>
<td>Commonly occurs on leaves, but may also affect leaf sheaths; infections of stems and heads are rare.</td>
<td>Commonly affects leaf blades, occasionally observed on heads when disease is very severe; infection of leaf sheaths or stems is rare.</td>
</tr>
<tr>
<td>Shape and distribution of lesions</td>
<td>Oval-shaped or elongated blister-like lesions scattered on affected tissues, lesions visible on both sides of leaf</td>
<td>Round or slightly elongated blister-like lesions scattered on affected tissues</td>
<td>Small, round, blister-like lesions that merge to form stripes</td>
</tr>
<tr>
<td>Lesion colour</td>
<td>Brown</td>
<td>Orange</td>
<td>Yellow</td>
</tr>
<tr>
<td>Degree of damage</td>
<td>Tearing of outer layers of plant tissue that is visible without magnification</td>
<td>Tearing of outer layers of plant tissue rare, visible with magnification</td>
<td>No tearing of outer layers of plant tissue</td>
</tr>
</tbody>
</table>

Table 1.1. Physical characteristics used to differentiate the wheat rust diseases (Wolf et al., 2010).

The three wheat rust pathogens have slightly different optimum temperature and moisture conditions for infection. Free water is an essential requirement for germination of all three pathogens and temperature conditions are important for germination, growth and disease development (Singh et al., 2002). Stem rust is a warm-temperature disease that develops optimally
between 15°C and 25°C; however the disease can occur at temperatures between 15°C and 35°C (Singh et al., 2002). Leaf rust develops optimally at temperatures between 15°C and 30°C, but the disease will progress at temperatures between 15°C and 40°C. Stripe rust requires temperatures less than 18°C, which are usually associated with higher elevations, northern latitudes and cooler climates (Marsalis and Goldberg, 2006). The optimum temperature for the development of this disease is 9-17°C, with disease progression ceasing at temperatures above 23°C (Marsalis and Goldberg, 2006, Singh et al., 2002, Murray et al., 2005). Because stripe rust is favoured by low temperatures, the disease occurs earlier in the growing season than leaf and stem rusts, and therefore, has the potential to cause more damage (Bushnell and Roelfs, 1985).

Under optimal conditions for disease development, infection occurs in 6-8 h and uredospores capable of causing secondary disease are produced on average 11-14 days after infection, although variations of these times can occur depending on the temperature and degree of susceptibility of the host (Murray et al., 2005, Marsalis and Goldberg, 2006). The capacity of uredospores to sequentially reinfect wheat crops in the same season leads to rapid epidemic spread of these diseases. Despite the fact that uredospores remain viable for only a limited length of time compared to other spore stages such as teliospores, they are spread extremely efficiently as they are produced in massive quantities and are easily dispersed by wind or on clothes, hands and equipment (Marsalis and Goldberg, 2006).

1.1.1. Economic impact of the wheat rusts

Globally, wheat is one of the most significant crops, with total production of 674.8 million tonnes in 2012 (FAOSTAT http://faostat3.fao.org/home/index.html). Wheat is grown on more land area than any other commercial crop and continues to be the most important food grain source for humans (Curtis, 2002). Wheat rusts are the most economically important group of wheat diseases, causing more than $5 billion in losses worldwide each year (Mwando et al., 2012). Losses due to leaf rust are primarily due to grain shrivelling and are usually less than 10 percent, but in highly susceptible genotypes losses can be more severe (Singh et al., 2002). Stem rust is the most devastating of the wheat rust diseases; losses of 100 percent can occur with susceptible cultivars (Singh et al., 2002). Stripe rust reduces the yield and quality of grain and forage. Seeds produced from crops damaged by this disease have low vigour and poor emergence after germination. Typically, yield losses caused by stripe rust in wheat producing areas range from 10-70% depending upon the susceptibility of cultivar, the earliness of the initial infection, the rate of disease development, and the duration of the disease
Chapter 1: Introduction to wheat stripe rust

However, losses can reach levels of up to 100% if the infection occurs very early in the season and the disease continues developing during the growing season (Afzal et al., 2007). Stripe rust has become increasingly important in the U.S. where epidemics since 2000 have caused yield losses of millions of dollars in spite large amounts of money spent on fungicides (Chen, 2005). In recent years, stripe rust has also caused severe damage to wheat in many other countries including South Africa (Pretorius, 2004), China (Wan et al., 2007) and Central Asia (Morgounov et al., 2004). In Australia, stripe rust is ranked as the second most important disease in terms of disease losses which is controlled by a combination of breeding and pesticides at a cost of AUD 900 million every year (Wellings, 2011, Wellings, 2007, Murray and Brennan, 2009). Stripe rust is likely to be the most destructive rust in cool and moist seasons, while stem and leaf rusts are likely to be more destructive in warm and moist seasons. However, this rule of thumb appears to be changing as in recent years new isolates of \textit{Pst} have shown significant adaptation to warm temperatures and enhanced aggressiveness, causing severe disease in previously unfavorable environments (Milus et al., 2009, Hovmøller et al., 2008).

1.2. STRIPE RUST: UNDERSTANDING THE PATHOGEN AND THE DISEASE IN WHEAT

1.2.1 Life cycle

Rust fungi have the most complicated life cycles of all fungi and can be categorized by how many types of spores are produced during the life cycle. For many rust fungi the life cycle is macrocyclic, in that it involves five spore stages, and heteroecious, in that it requires two unrelated hosts for its completion. \textit{Pt} (leaf rust) and \textit{Pgt} (stem rust) are both macrocyclic and heteroecious (Kolmer et al., 2009). \textit{Pst} was traditionally classified as a hemicyclic and autoecious pathogen, which means it only has three spore stages (uredospores, teliospores, and basidiospores), and produce spores only on wheat or very closely related host plants (Carver, 2009). Uredospores were the only known source of inoculum, and basidiospores and teliospores were thought to be not significant in the \textit{Pst} life cycle. Very recently, Jin et al. (2010) demonstrated that \textit{Pst} can also use \textit{Berberis} spp. as alternate hosts under laboratory conditions, which allows \textit{Pst} to complete the macrocyclic and heterocuous lifecycle (Figure 1.2).
Chapter 1: Introduction to wheat stripe rust

Figure 1.2. Life cycle of *Puccinia striiformis*. The asexual cycle of *Pst* on wheat occurs through uredospores (a), and later in the season the black two-celled teliospores appear (b) which germinate to produce a basidium, which in turn produces basidiospores (c). Basidiospores infect the alternative host barberry, where the pycnial (d) and aecial (e) stages appear on infected leaves. Aeciospores are a potential source of inoculum for wheat wherever the alternate host is found. The different spore stages can be haploid (n), diploid (2n) and dikaryotic (n+n). Modified from (Hovmoller et al., 2011)

*Pst* produces single-celled, dikaryotic (n+n) uredospores on leaves, leaf sheaths, and glumes of wheat. The uredospore stage is responsible for large disease outbreaks through the asexual multiplication and spread of uredospores as airborne inocula, leading to repeated infections in the same growing season. Later in the season, two-celled black teliospores may form along uredia (the pustule-like structure that produces uredospores). Teliospores are diploid (2n) and as they germinate, the nuclei undergo karyogamy and subsequently meiosis, giving origin to a four-celled basidium with haploid basidiospores (n). Basidiospores can infect the alternate host barberry and if they do, result in pycnia infections on the upper side of the leaf followed by aecial development on the lower side of the leaf (Jin et al., 2010). The completion of the sexual life cycle occurs when
aesciospores infect wheat leaves, which ultimately leads to production of the typical yellow rust uredia arranged into stripes (Jin et al., 2010). The existence of an alternative host potentially has an important role for the disease, not only because it makes it possible for recombination of virulence genes, but also because it allows the pathogen to survive between cropping seasons (Jin, 2011). Pst populations in countries such as China and Pakistan have complex pathogenic structures that apparently cannot be accounted for by single-step mutations to virulence (Enjalbert, 2010, Duan et al., 2010, Bahri et al., 2011). The report of Berberis spp. as the sexual stage host for Pst and the predicted distribution of these hosts in these regions suggests that the sexual cycle of the pathogen might influence the emergence of new pathotypes in nature (Park and Wellings, 2012). However, the extent and significance of sexual recombination in Pst in nature remains to be demonstrated.

1.2.2. Infection of wheat by Pst uredospores

The infection process in wheat starts once uredospores from local or distant regions are deposited on the leaf surface. A minimum of three hours of continuous moisture on the plant surface and temperature of 9-12°C are required for uredospores to imbibe water, swell and produce a germ tube which continues to elongate until it comes into contact with a stomate (Carver, 2009). The host leaf topography is very important for correct germ tube orientation; rusts inoculated onto non-host plants are unable to locate stomata, or may not even germinate properly (Wynn and Staples, 1981). Once a stomate is found, the germ tube tip enters the leaf interior via a stomatal opening, then differentiates into a substomatal vesicle (Moldenhauer et al., 2006). This vesicle typically forms 2–3 primary infection hyphae which are delimited by a septum, and forms a haustorial mother cell (HMC) after contacting a mesophyll cell wall. The HMCs penetrate the plant cell walls and form haustoria, which expand between the plant cell wall and the plasma membrane, thus remaining outside the host cytoplasm (Hovmoller et al., 2011).

The haustorium is probably the most important structure for understanding the rust host-pathogen interaction (Catanzariti et al., 2007, Voegele and Mendgen, 2011) and its function and development will be described in more detail later. After the establishment of the first haustorium, secondary infectious hyphae start to develop from the primary hyphae and grow systemically in the intracellular spaces between mesophyll cells to form an extensive hyphal network. Days later, under the leaf epidermis, the invasive hyphae form sporogenous basal cells in the uredia, and thousands of uredospores bud from that structure and erupt through the epidermis, from which point they can be dispersed to start a new disease cycle on wheat. Under optimal temperatures, the process from uredospore landing to release of new uredospores takes 10-13 days. The fungus continues to
produce new uredia and uredospores in stripe patterns further up and down the leaf from the initial site of infection (Carver, 2009). This process is illustrated in Figure 1.3.

![Figure 1.3. The infection process of rust pathogens on wheat.](image)

Figure 1.3. The infection process of rust pathogens on wheat. Modified from (Voegele et al., 2009) and (Kang et al., 2001). The arrows indicate fungal and plant structures. Plant cell wall (PCW), plant cytoplasm (PC), extrahaustorial matrix (EHMx), extrahaustorial membrane (EHM).

### 1.3. BIOTROPHIC INTERFACES

Plant-associated fungi and oomycetes range from obligate biotrophs, which require a living host to complete their life cycle and cannot be cultured in vitro, to necrotrophs which promote the destruction of host cells to feed on their contents and can freely exist outside of plants (Yi and Valent, 2013). Hemibiotrophs have an intermediate lifestyle, with an initial biotrophic phase and a subsequent necrotrophic stage. Obligate biotrophs include the powdery mildews (Ascomycota), the rusts (Basidiomycota), and the downy mildews and white rusts (Oomycota). While they each colonize their host in different ways, all successfully establish a long-lasting intimate relationship with their hosts. The formation of highly specialized hyphae with terminal feeding structures such as haustoria serves as the main interface between the pathogen and its host (Eichmann and Huckelhoven, 2008, Yi and Valent, 2013). Looking at the broad phylogenetic spectrum of haustoria-forming organisms, it appears likely that these structures have arisen more than once in the course of evolution (Voegele and Mendgen, 2003). Similar membrane-bound structures have evolved in other organisms, for example the arbuscules produced by arbuscular mycorrhiza (obligate fungal symbionts of land plants) or modified roots (also called haustoria) produced by parasitic plants,
suggesting that such architecture represents a particularly successful adaptation of these organisms to interact with their respective host plants to solve a common problem (Yoder, 1999, Voegele and Mendgen, 2003).

1.3.1. The haustorial complex

The haustorium has been of major scientific interest ever since its discovery more than 150 years ago (von Mohl, 1853). Rust fungi derive energy from living host cells to sustain growth and reproduction, and haustoria play the primary role in nutrient uptake from the host (Voegele and Mendgen, 2003, Voegele and Mendgen, 2011). Haustoria are only formed in planta, making them extremely difficult to study. Although axenic cultures have been established for some biotrophs (Mendgen and Hahn, 2002), most of the economically important biotrophic fungi remain nonculturable, at least not to a point equivalent to growth and reproduction in the biotrophic mode. Additionally, the lack of genetic transformation systems for biotrophic fungi excludes these organisms from the use of classical molecular tools to study them. Haustoria can be produced by either mono- or dikaryotic mycelia of rust fungi. However, monokaryotic haustoria present as intracellular extensions of intercellular hyphae with no significant morphological specialization (Gold and Mendgen, 1991). On the other hand, dikaryotic rust haustoria develop from specialized haustorial mother cells which are external to the host cell wall, and when mature, consist of two distinct regions with significant differentiation, the haustorial neck and the haustorial body. The haustorial neck and body remain separated from the host cytoplasm of the invaded cell by a region known as the extrahaustorial matrix (EHMx) and the extrahaustorial membrane (EHM) (Figure 1.3). Thus, haustoria are not truly intracellular structures, and always remain outside the plant plasma membranes. Several lines of evidence suggest that the EHM is functionally distinct from the plant plasma membrane. These include the lack of the typical H+/ATPase activity of plant plasma membranes (Gay et al., 1987, Manners, 1989, Baka et al., 1995), the lack of intramembrane particles (Harder and Chong, 1991) and the absence of certain proteins that are specifically associated with plant plasma membranes (Koh et al., 2005, Micali et al., 2011). Green fluorescent protein (GFP) tagging was used to show that eight plasma membrane marker proteins are all absent from the EHM of Erysiphe cichoracearum, leading to suggestions that this membrane is newly synthesized with the formation of the haustorium (Koh et al., 2005). Freeze fracture studies have shown the EHM typically has corrugations and protuberances that increase with the age of the haustoria (Harder and Chong, 1991). It is thought that convolutions in invaginated membranes act to increase the surface area, thus aiding nutrient acquisition by the fungus (Manners and Gay, 1982). Additionally, the lack of
ATPase activity could mean that there would be little or no control over solute fluxes from the host cell, favoring uptake of nutrients by the fungus (Voegele and Mendgen, 2003). This model is supported by the observation that in uninfected soybean leaf discs, sucrose efflux is promoted by inhibition of plasma membrane ATPase activity and by elevation of the extracellular pH (Secor, 1987). The plasma membrane stain periodic acid-chromic acid-phosphotungstic acid (PACP) often produces a notably reduced intensity at the EHM compared to the plant plasma membrane, probably due to a reduction in glycoprotein content (Woods and Gay, 1983). Freeze-fracture transmission electron microscopy following filipin treatment (an affinity stain for sterols) revealed an absence of granular filipin-sterol complexes on the EHM of the rust fungi Puccinia coronata and Uromyces appendiculatus (Harder and Mendgen, 1982). This suggests that the EHM contains less sterol than normal plasma membranes, which could affect both its fluidity and permeability. One of the few known EHM-specific proteins is a glycoprotein identified in the EHM of the powdery mildew Erysiphe pisi (Roberts et al., 1993). This glycoprotein was found to be present from the earliest stages of haustorium formation but was not detectable in older haustoria, suggesting that the EHM changes its composition during haustorial development. Studies on the Arabidopsis-powdery mildew interaction found that the plant resistance protein RPW8.2, which is associated with broad-spectrum resistance to powdery mildew, is targeted specifically to the EHM of Golovinomyces cichoracearum but not to the host plasma membrane (Wang et al., 2009a). RPW8.2 was also detected in vesicles in the plant cytoplasm adjoining haustoria, suggesting the existence of a dedicated secretory pathway targeting plant proteins to the EHM (Wang et al., 2009a). The exact composition of the EHM remains one of the major questions in biotrophy. Its identity, the proteins that are unique to this biotroph interface, the mechanisms that specifically target them to this structure, and how its content is manipulated during development, remain elusive.

The presence of an EHMx, surrounding the haustorium and separating it from the plant cell, has been observed in almost all haustoria forming species (Perfect and Green, 2001). In rust fungi, this structure appears to be an amorphous mixture of components, mainly carbohydrates and proteins (Harder and Chong, 1991). It is currently not possible to isolate the EHMx for direct analysis of its composition, so most information about the nature of this structure comes from in situ cytochemistry studies. Most evidence points to the conclusion that the matrix consists partly of fungal and partly of plant-derived materials (Harder and Chong, 1991), but the degree of contribution from each source appears to be specific to the particular host-pathogen interaction (Perfect and Green, 2001). Many compounds that are usual components of plant cell walls have been detected in the rust EHMx, including polysaccharides such as pectins and xyloglucans, and
glycoproteins including arabinogalactan proteins, hydroxyproline-rich proteins and threonine-hydroxyproline-rich proteins (Stark-Urnau and Mendgen, 1995, Hippe-Sanwald et al., 1994). Conversely, plant cell wall components are generally not detectable in extrahaustorial matrices of fungal powdery mildews (Celio et al., 2004). One possibility is that secretion of the host components is a resistance response, and the pathogen suppresses the synthesis and secretion of these cell wall components into the interface zone (Green et al., 2002). Alternatively, these components might be rapidly degraded by fungal hydrolytic enzymes once they are secreted into the EHMx, and perhaps could become a source of nutrients for the pathogen (Green et al., 2002). The EHMx generally does not contain fungal wall polysaccharides such as chitin, but immunolabelling has demonstrated the presence of fungal proteins. Proteins of rust fungi such as an elicitor glycoprotein of Pgt (Marticke et al., 1998) and the invertase Uf-INVI and glycoprotein Uf-RTP1 of U. fabae (Voegele et al., 2006, Kemen et al., 2005) have been detected in the EHMx. Indirect evidence supporting the fungal contribution to the formation of the EHMx comes from mutants of the maize smut fungus Ustilago maydis. Biotrophic intracellular hyphae (a structure similar to the haustorium but with the capacity to pass from one cell to the next) defective in an α-glucosidase was arrested in growth within the leaf epidermis, and the interface between mutant hyphae and the plant plasma membrane was altered compared with the interface of wild-type hyphae (Schirawski et al., 2005). Given that the role of α-glucosidase is to process N-linked glycoproteins in the endoplasmic reticulum, this observation is consistent with the idea that proper glycosylation of fungal cell wall or matrix proteins is essential for establishment of the functional biotrophic interface in this pathosystem (Schirawski et al., 2005). The initial biotrophic phase of hemibiotrophs like Colletotrichum lindemuthianum is also characterized by the presence of an interfacial matrix separating host and parasite plasma membranes (Perfect and Green, 2001). A gene cloned from C. lindemuthianum encodes a protein specifically localised to this matrix which resembles plant cell wall proline- and hydroxyproline-rich proteins, suggesting that the fungus might use mimicry to achieve biotrophy (Perfect and Green, 2001, Perfect et al., 1998). The EHMx is believed to function as a key ‘trading place’ for the uptake of nutrients into the pathogen and the export of virulence effector molecules into the host cell (discussed in a later section). Electron microscopy studies on the host-pathogen interface between daylily leaf cells infected by Puccinia hemerocallidinis revealed the presence of large aggregations of tubular cytoplasmic elements near haustoria in infected host cells (Mims et al., 2002). Many of these elements were continuous with the EHM and electron-dense deposits present in the EHMx extended into these elements. However, it remains to be shown if there is any kind of trafficking linked to these structures. Micali et al. (2011) detected vesicles delimited by a lipid bilayer in the EHMx of Golovinomyces orontii haustoria in Arabidopsis. Whether the observed vesicles represent exosomes
Chapter 1: Introduction to wheat stripe rust

derived from the fusion of fungal multivesicular bodies with the haustorial plasma membrane, or plant vesicles produced by budding-off from the EHM is still unknown. Despite its unresolved origin, it appears likely that the EHM plays critical roles in the maintenance of the biotrophic lifestyle and serves as a space for exchange of compounds between the pathogen and the host.

The haustorium is connected to the rest of the fungal body through a narrow junction referred to as the haustorial neck. A common feature surrounding the neck is a structure often referred to as the “neck ring” or “neckband”. The neckband binds to the host plasma membrane and demarcates the point at which it becomes the EHM and the EHMx begins (Harder and Chong, 1991). The neckband appears to prevent the release of host nutrients from the EHMx into the host apoplast (Panstruga, 2003, Voegele and Mendgen, 2003). Koh et al. (2005) found that the neck region marked the site where the EHM differentiates from the host plasma membrane, and they suggested models involving the neck band as a key component in formation of the EHM. In most haustoria-forming fungi, the neck region is characterized by the presence of some electron-dense material, seemingly joining the two plasma membranes of host and parasite, and it is often referred to as the collar (Harder and Chong, 1984). Ultrastructural studies on the rust fungus *Frommeëla mexicana* var. *indicae* equated collars with the cell wall ingrowths found in plant transfer cells (Mims et al., 2001). Transfer cells are types of parenchyma cells that possess distinctive labyrinthine wall ingrowths that greatly increase the surface area; they appear to facilitate solute transfer over short distances and are common at sites where elevated nutrient flow is necessary for plant development (McCurdy et al., 2008). Therefore cell wall ingrowth might serve as a host cell response to compensate for solutes lost from the cell to the haustorium. Nevertheless, the function and composition of the collar remain to be elucidated.

1.3.2. Nutrient uptake by rust haustoria

As early as 1863 it was proposed that one of the possible functions for haustoria would be the uptake of nutrients from the host (de Bari, 1863). The development of procedures to purify haustoria of rust fungi from infected plant tissue was a major technical advance that made the molecular analysis of haustoria possible (Hahn and Mendgen, 1992). Hahn and Mendgen (1997) were able to construct a cDNA library from the bean rust fungus *U. fabae* which was an extraordinarily rich source of in planta-induced fungal genes (PIGs) (Hahn and Mendgen, 1997). Several of these genes showed sequence similarity to genes involved in nutrient acquisition and many of them were subsequently investigated in greater detail. The first of these genes is the hexose
transporter \textit{HXT1}, which was found to localize exclusively at the haustoria plasma membrane, giving the first compelling evidence for involvement of haustoria in nutrient uptake (Voegele et al., 2001). Heterologous expression of \textit{HXT1} in \textit{Saccharomyces cerevisiae} and \textit{Xenopus laevis} oocytes showed that \textit{HXT1}p has substrate specificity for D-glucose and D-fructose and indicated a proton symport mechanism (Voegele et al., 2001). Years later, the invertase gene \textit{Uf-INV1} in the same pathogen was identified, and the encoded protein was found to be secreted into the EHMx (Voegele et al., 2006). These studies provided the first proof that haustoria are critical for sugar uptake. Among the \textit{PIG}s, three genes encoding putative transporters for amino acids were identified and were subsequently renamed \textit{AAT1}, \textit{AAT2} and \textit{AAT3} (Struck et al., 2002, Struck et al., 2004, Mendgen et al., 2000). \textit{AAT1} and \textit{ATT3} were shown to function as proton-dependent transporters with preference for histidine/lysine and leucine/methionine/cysteine respectively, expressed predominantly but not exclusively in haustoria (Struck et al., 2002, Struck et al., 2004, Mendgen and Hahn, 2002). \textit{AAT2} localized exclusively to the haustorial plasma membrane, but its specific substrate has not been identified (Mendgen et al. 2000). These observations suggest that amino acid uptake in \textit{U. fabae} is not limited to haustoria, but the transporters characterized are energized by the proton-motive force, and show a preference for amino acids present in low abundance in infected leaves (Struck et al. 2004). More recently, sequencing of the genomes of the rust pathogens \textit{Pgt} and \textit{Melampsora larici-populina}, complemented with transcriptomic data, allowed the identification of homologs to \textit{HXT1, ATT1, AAT2} and \textit{AAT3} and several other putative sugar and amino acid transporters, many expressed \textit{in planta} (Duplessis et al., 2011). Furthermore, genes encoding putative ion transporters and auxin transporters were found in the genomes of these pathogens, several of them expressed \textit{in planta}, giving additional support to the role of haustoria in nutrient uptake from the host.

\textbf{1.3.3. Additional roles of haustoria}

Biosynthesis of important molecules is an additional role attributed to haustoria, which presumably is facilitated by the availability of resources taken up from the host. Thiamine is an essential cofactor required for the activity of key enzymes involved in carbon metabolism (Sohn et al., 2000). Massive expression of the thiamine biosynthesis-related genes \textit{THI1} and \textit{THI2} was found in the haustoria of \textit{U. fabae} (Sohn et al., 2000). Complementation studies in a heterologous system confirmed the role of these genes in the synthesis of the pyrimidine moiety (\textit{THI1}p) and the thiazol moiety (\textit{THI2}p) (Sohn et al., 2000) of thiamine. Immunofluorescence microscopy on rust-infected leaves detected high concentrations of \textit{THI1}p in haustoria, in a lower concentration in the basal hyphae of the uredia.
and in very low levels in uredospores (Sohn et al., 2000). These results suggest that the haustorium is probably the major site for synthesis of thiamine.

Another haustorially-expressed gene, PIG8, encodes a NADP+-dependent mannitol dehydrogenase, subsequently designated as MAD1 (Voegele et al., 2005). Biochemical studies of this enzyme revealed that MAD1p would be likely to synthesize mannitol from fructose (forward reaction) in haustoria, while in germinating spores the reaction would proceed backwards toward the formation of fructose from mannitol (reverse reaction) (Voegele et al., 2005). Mannitol is a secondary metabolite typical of fungi, and can have a role in tolerance to environmental stresses or act as a mobile carbon storage compound (Solomon et al., 2007, Voegele and Mendgen, 2003). Large amounts of mannitol have been detected in uredospores of *U. fabae*, which disappear rapidly upon germination (Voegele et al., 2005). MAD1 transcripts were detected only in haustoria, while MAD1p could be found in the lumen of haustoria and in uredospores (Voegele et al., 2005). In uredospores, where there is low fructose and high mannitol concentrations, MAD1p appears to be important in making fructose available for quick entry into glycolysis. Thus, MAD1p appears to be responsible for the formation of mannitol from D-fructose in haustoria, and for the mobilization of mannitol for metabolism in germinating uredospores (Voegele et al. 2005)(Voegele and Mendgen, 2011). D-arabitol dehydrogenase 1 (ARD1p) was a novel enzyme detected in *U. fabae*, which also localized in haustoria (Link et al., 2005). This enzyme acts on from xylulose and ribulose (NADP+-dependent) to produce arabinol, a polyol that have been found in infected leaves and ungerminated uredospores. However, since no ARD1p was detected in spores, utilization of D-arabitol in germinating spores has to proceed via a different enzymatic pathway (Link et al., 2005). Haustoria appear to be a place of major metabolic activity, important for example in the production of polyols. However, how polyols are mobilized from haustoria, trafficked through the mycelium and deposited into the uredospores, remains unknown (Voegele and Mendgen, 2011).

### 1.4. BIOTROPHIC EFFECTORS

Currently, a very active field of research is in the analysis of proteins and other molecules secreted by bacteria, fungi, oomycetes, and nematodes, into their hosts to manipulate various aspects of the interaction in favour of the pathogen (Hogenhout et al., 2009, Misas-Villamil and van der Hoorn, 2008, Kamoun, 2006, Kamoun, 2007). The functional characterization of these secreted molecules is considered to be fundamental to understand the pathogens’ strategies to colonize their hosts and cause disease (Hogenhout et al., 2009). These proteins and molecules include toxins, degradative
enzymes, and proteins that act as virulence factors, which are now combined under the general term effectors (Kamoun, 2006). The use of this term became widely used by workers studying phytopathogenic bacteria after the discovery that they encode specialized machinery, termed the type three secretion system (TTSS), to deliver proteins inside host cells (Van Gijsegem et al., 1993, Alfano and Collmer, 1996, Alfano and Collmer, 2004). The TTSS acts like a “molecular syringe” that contacts host cells, and effectors are translocated through it into the plant cytoplasm (Jin et al., 2003). The TTSS proteins can be grouped into three categories: 1. Structural proteins which build the base (anchored to the bacterial membrane) and the needle, 2. Effector proteins that get secreted into the host cell, and 3. Chaperones that bind effectors in the bacterial cytoplasm to stabilize and prevent association before delivery and to direct them towards the needle complex (Akeda and Galan, 2005, Alfano and Collmer, 2004, Marlovits et al., 2004). Contact of the TTSS with a host cell initiates secretion, but external cues such as temperature, pH, osmolarity and oxygen levels, can also induce activation. Several transcription factors that regulate the expression of T3SS genes are known.

The delivery of effectors into the host cell cytoplasm requires an additional set of three proteins, called translocators, which are themselves exported by the needle. These proteins are involved in the formation of a translocation pore, the translocon, in the host cell membrane, through which other effectors may enter (Alfano and Collmer, 2004). Several of the best-studied TTSS effectors were defined as *avirulence* (*Avr*) genes due to their abilities to trigger the hypersensitive response (HR) on resistant plants. In these situations, the *Avr* gene is recognized by a matching *Resistance* (*R*) gene in the host, a phenomenon called gene-for-gene resistance (Keen, 1990, Staskawicz et al., 1984). The key observation over a number of years by several laboratories is that some *Avr* proteins actually confer heightened virulence activity on compatible hosts (i.e. hosts that lack the matching *R* gene) (Luderer and Joosten, 2001). Thus, effectors are generally associated with positive roles in establishing infection (Alfano and Collmer, 2004)). Some bacterial effectors are targeted to the host nucleus and may act as transcription factors (Lahaye and Bonas, 2001, Deslandes et al., 2003), while others have protease activities to cleave specific host proteins (Orth et al., 2000, Axtell and Staskawicz, 2003, Coaker et al., 2005), and various other activities are known (Block et al., 2008).

Contrasting with bacteria, little is known about the identity, release, uptake, and function of effectors of biotrophic fungal pathogens. Currently, a great deal of effort is being put into the identification of effector genes, targets and how effector–target interactions promote the pathogenesis of plants.

During the last decade, some proteins secreted by these pathogens have been characterized, and they can be grouped roughly into extracellular effectors that are secreted into the apoplast or xylem
of their host plants, and cytoplasmic effectors that are translocated into host cells (De Wit et al., 2009). Despite the low degree of sequence conservation amongst fungal effectors, many code for small secreted proteins (containing an N-terminal signal peptide (SP) for secretion), some of which are translocated into host cells by an as yet unknown mechanism(s). Another feature of many fungal effectors is the presence of multiple cysteine residues that might be involved in disulphide bridge formation, providing protein stability (Stergiopoulos and de Wit, 2009). However, as yet there are not sufficient defined criteria such that fungal effectors can be recognized by molecular characteristics alone, making the discovery and characterization of these pathogenicity factors more challenging.

1.4.1. Extracellular effectors

Extracellular effectors act in the extracellular space at the plant-microbe interface, where the early plant innate immune response is commonly suppressed to help the pathogen to adapt to the pathogenic niche (Hogenhout et al., 2009, Misas-Villamil and van der Hoorn, 2008, Kamoun, 2006). One good example is the hemibiotrophic fungal pathogen of tomato, *Cladosporium fulvum*, which remains exclusively in the apoplast during the infection cycle and does not form haustoria or haustoria-like structures (Rivas and Thomas, 2005). To date, four *Avr* genes have been cloned from this pathogen (*Avr2, Avr9, Avr4, and Avr4E*), and all encode small cysteine-rich proteins that are secreted during infection. Their recognition in tomato is mediated by the cognate R proteins Cf-2, Cf-4, Cf-4E and Cf-9 respectively (De Wit et al., 1997, Thomma et al., 2005, Stergiopoulos and de Wit, 2009). *Avr2* is a cysteine protease inhibitor that protects the fungus from the deleterious effect of the tomato cysteine proteases Rcr3, PIP1 and TDI-65 (Rooney, 2005, Shabab et al., 2008). Similarly, *AVR4* binds to chitin in fungal cell walls to protect it from tomato host chitinases and to prevent chitin fragments from triggering plant immunity (van den Burg et al., 2006, Libault et al., 2007). Additional extracellular proteins (ECPs), namely ECP1, ECP2, ECP4 and ECP5, have been characterized from *C. fulvum*, and trigger the HR in tomato cultivars that carry a cognate *Cf-Ecp* R gene (De Kock et al., 2005). ECP6 and ECP7 have also been identified, but no R genes capable of recognising either of these two effectors has yet been reported (Bolton et al., 2008). ECPs are abundantly secreted by all strains of *C. fulvum* during infection, and possess an even number of cysteine residues that are most probably involved in intramolecular disulphide bridges (De Wit et al., 2009). All *AVRs* and ECPs are assumed to be virulence factors, although this has been demonstrated only for *Ecp6, Ecp1 and Ecp2*, as silencing or disruption of these genes compromises the virulence of *C. fulvum* on tomato (Bolton et al., 2008, Thomma et al., 2005, Stergiopoulos and de Wit, 2009).
Oomycetes such as *Phytophthora* spp. which have similar infection strategies to fungi but belong to the Stramenopile kingdom also secrete apoplastic effectors (Birch et al., 2006, Damasceno et al., 2008, Kamoun, 2006). A well-known example is the glucanase inhibitor protein 1 (GIP1) of *P. sojae* which differentially targets β-1,3 endoglucanases of various *Glycine* (bean) species (Bishop et al., 2005). *Phytophthora infestans* secretes cysteine protease inhibitors, such as EPIC2B, which inhibits PIP1 and other apoplastic cysteine proteases of tomato (Tian et al., 2007). Effectors EPI1 and EPI10 are multidomain serine protease inhibitors of the Kazal protein family that are secreted by *P. infestans*, which inhibit efficiently the tomato pathogenesis-related subtilisin-like serine protease protein P69B, thought to function in defence (Tian et al., 2005). In the biotrophic fungus *U. maydis*, a novel secreted protein termed PEP1 was found to be secreted into the apoplast and to be essential for penetration (Doehlemann et al., 2009). *pep1* mutants penetrate the plant cell wall and invaginate the plant plasma membrane, but any further fungal development is arrested. Further investigation of this effector showed that the protein accumulates in the interface around the intracellular biotrophic hyphae and at sites of cell-penetrations where it blocks the oxidative burst via a direct inhibition of host peroxidases (Hemetsberger et al., 2012), giving the first example of a fungal effector directly interfering with the reactive oxygen species (ROS) -generating system of the host plant. A deep transcriptome sequencing approach on the hemibiotrophic fungus *Colletotrichum higginsianum* allowed the identification of candidate effector genes, and their consecutive waves of expression were associated with key developmental transitions, indicating that distinct suites of effectors are deployed at each infection stage (Kleemann et al., 2012). The focal secretion of effector proteins at the interface formed between host and pathogen in the very early stages of infection was demonstrated using fluorescent protein tagging of effectors and transmission electron microscopy-immunogold labelling (Kleemann et al., 2012). Ultrastructural analyses revealed that host-derived cell wall material is deposited beneath appressoria before any visible penetration damage to the cuticle, so it was proposed that the fungus deploys early-expressed effectors to counteract pre-invasion host defences and to prepare the host cell for colonization. All this reinforces the idea that apoplastic effectors are crucial for the establishment of the pathogen and the initiation of colonization, yet constitute an almost unexplored area in the rust fungi.

### 1.4.2. Cytoplasmic effectors

A substantial portion of the effector proteins secreted by biotrophic/hemibiotrophic pathogens are thought to act inside the host cell. The first indications that effectors are translocated into the host cell cytoplasm were derived from the fact that many of them are recognized as *Avr* determinants by
plant R proteins with either proven or assumed cytoplasmic localization (Dodds and Rathjen, 2010, O’Connell and Panstruga, 2006). Supporting this idea, the first fungal effectors were identified in attempts to clone Avr proteins recognized by host R proteins, and it became apparent that many were recognized within the host cell cytoplasm (Dodds et al., 2004, Jia et al., 2000). Fungal effectors presumed to act intracellularly appear to be synthesized in haustoria or intracellular hyphae, transported through a secretory pathway and finally secreted into the interfacial matrix via exocytosis (Giraldo et al., 2013, O’Connell and Panstruga, 2006). This idea is supported by the fact that transcripts encoding certain effectors are enriched in haustoria (Catanzariti et al., 2006), and that many (but not all) of the identified intracellular effectors carry prototypical SP signals (O’Connell and Panstruga, 2006). However, the mechanism by which this transfer is achieved remains elusive (Ellis et al., 2006). Studies on oomycetes indicate that secreted effector proteins carry an N-terminal RxLR motif (Arg-x-Leu-Arg, where x indicates any amino acid) together with a downstream dEER (Asp-Glu-Glu-Arg) motif, which together are required for pathogen-independent internalization (Morgan and Kamoun, 2007, Whisson et al., 2007, Jiang et al., 2008, Haas et al., 2009, Tyler et al., 2013). The RxLR motif of several oomycete effectors binds to phosphatidyl inositol phosphates (PIPs) on the outer surface of the plant plasma membrane and they possibly enter the cell through lipid raft-mediated endocytosis (Kale et al., 2010). Similar pathogen-independent uptake was demonstrated for effectors from the biotrophic flax rust fungus, Melampsora lini (Rafiqi et al., 2010). In planta transient expression of the flax rust effector proteins AvrM and AvrL567 including their signal peptides (which are functional in plants) resulted in intracellular accumulation of these proteins without requiring a pathogen-encoded transport mechanism (Rafiqi et al., 2010). However, results from lipid-binding assays on flax rust effectors do not support the phospholipid-mediated uptake model of oomycete effectors (Gan et al., 2010). In a search for host targeting signals in fungi, Godfrey et al. (2010) showed that some small secreted proteins from B. graminis and Puccinia spp. possess a conserved amino acid motif Y/F/WxC. No function has yet been ascribed to this motif (Godfrey et al., 2010). However, systematic searches for the motif in predicted effectors from the poplar leaf rust and the wheat stem rust genomes showed that the motifs were indeed present in several of these genes, but not restricted to the N-terminal region as in B. graminis (Duplessis et al., 2011). Additionally, these motifs were also present in non-secreted proteins related to zinc binding and nucleic acid binding suggesting that they are not necessarily related to or diagnostic of effector function as the RxLR motif apparently is (Duplessis et al., 2011). Very recently, Giraldo et al. (2013) provided evidence that the hemibiotrophic rice pathogen Magnaporthe oryzae possesses two separate and specialized routes by which it secretes effector proteins during biotrophic invasion of rice. While apoplastic effectors accumulate extracellularly at the host–pathogen interface and are
actively secreted via the conventional endoplasmic reticulum-to-Golgi secretory process, cytoplasmic effectors destined for delivery inside rice cells are secreted by a different pathway involving the exocyst complex (Giraldo et al., 2013). Using a combination of pharmacological and genetic analyses, they showed that cytoplasmic effectors require the exocyst components Exo70 and Sec5 and the Sso1 t-SNARE for efficient secretion. How these processes function and their evolutionary origins could be pivotal to our understanding of the distinct secretion systems in fungal pathogens which facilitate tissue invasion and regulation of virulence activities.

1.4.3. Effector identification and validation

In the past, the classic strategy to identify and validate effector genes was to determine whether they confer avirulence to a previously virulent pathogen strain on a host containing a particular R gene (Chisholm et al., 2006, Alfano, 2009). A good example is Avr-Pita of M. oryzae, the gene for which specifically confers avirulence toward rice cultivars that contain the resistance gene Pi-ta (Orbach et al., 2000). Pwl1 and Pwl2 are two additional Avr genes of M. oryzae identified with the same principle, as these genes are responsible for the non-pathogenicity of rice pathogens against Eragrostis curvula, weeping lovegrass (Kang et al., 1995, Sweigard et al., 1995). This R-Avr recognition approach was successfully performed in M. lini where the first flax rust effector identified was AvrL567, a protein recognized by the L5, L6 and L7 R proteins (Dodds et al., 2004).

More recently, advances via next-generation sequencing (discussed later) and whole-genome sequencing projects of fungal pathogens has proved to be a valuable tool that allows the identification of putative effector genes (O'Connell and Panstruga, 2006). The presence of a secretion signal peptide is considered a universal prerequisite to identify effectors, although there are a few exceptions (Ridout et al., 2006). Additional criteria such as small size, the presence of multiple cysteine residues, and localization of the gene in highly dynamic genomic regions, have been proposed to facilitate their identification (Hogenhout et al., 2009, Stergiopoulos and de Wit, 2009). However, there is not yet sufficient evidence to consider these criteria mandatory features of effector genes. Genome sequencing projects have allowed the identification of large sets of candidate effector genes in different biotrophic and hemibiotrophic plant pathogens such as Pgt and M. larici-populina (Duplessis et al., 2011), B. graminis (Spanu et al., 2010), U. maydis (Kamper et al., 2006), M. lini (Nemri et al., 2014), M. oryzae (Dean et al., 2005), and Colletotrichum spp. (Gan et al., 2013, O'Connell et al., 2012). The difficulty with this approach is that it requires a great deal of bioinformatics resources and expertise to assemble large, complex and highly repetitive genomes, as
is the case for many biotrophic fungi (Martin and Grigoriev, 2013). Transcriptome sequencing projects have proven to be extremely useful in the absence of an assembled genome sequence, since in addition to allowing prediction of effector genes, they usually provide information about the relative expression of the predicted genes (Wessling et al., 2012). However, the latter approach can be extremely challenging when applied to obligate biotrophs, which cannot be cultivated in vitro.

Typically, secreted effector proteins have low or no sequence homology to known proteins or protein sequence motifs, hindering identification of their functions. Ideally, gene overexpression, disruption or silencing assays would give clues about the roles of these genes during infection. However, for the majority of biotrophic fungal pathogens, transient or stable genetic transformation methods are yet to be developed (Voegele et al., 2009), and in those transformable hemibiotrophic fungi, roles for given genes can be hampered by apparent functional redundancy which can mask any obvious phenotype when individual genes are deleted or silenced (Muller et al., 2008, Mosquera et al., 2009). Therefore, the Avr-R recognition principle continues to be used to validate in silico predictions in various effector hunting strategies. One of the best known examples of this is with flax rust. The screening of a flax rust haustorial cDNA library led to the identification of 21 genes that encode secreted proteins, several of which induced HR with the appropriate $R$-gene specificity when expressed in flax, confirming their avirulence activities (Catanzariti et al., 2006). Other pathogens effector proteins have been validated by the Avr-R recognition approach, for example AVR$_{a10}$ and AVR$_{a1}$ in *Blumeria graminis* f sp *hordei* (Ridout et al., 2006) and Avr22 in *Pgt* (Updahyaya et al. unpublished).

Another important challenge in the validation/characterization of effectors predicted from genomic or transcriptomic studies is related to the large numbers of candidate genes to test. Additional criteria to narrow the number of candidates are necessary to facilitate their functional characterization. Similarity to known proteins (including similarities to Avr proteins), or exclusive gene expression during infection, have been used as such criteria. For example, Mueller et al. (2008) first searched the *Ustilago maydis* genome for putative secretion signal-containing genes and then narrowed the number of candidates down by searching for domains and motifs needed for proteins to function in the apoplast or inside eukaryotic cells. They identified putative effectors possessing nuclear localization signals, zinc finger domains, or RING domains, which suggested these proteins could act inside plant cells (Mueller et al., 2008). In *M. oryzae*, expression of mRNAs derived from invasive hyphae in the primary invaded host cells was compared with axenically-grown mycelium mRNAs using oligoarrays (Mosquera et al., 2009). Using a 10-fold differential expression threshold,
58 candidate effector genes were identified, and four of these were confirmed to be fungal biotrophy-associated secreted proteins, secreted into rice cells in distinct patterns in compatible interactions (Mosquera et al., 2009). Moreover, a putatively secreted chorismate mutase found in the *U. maydis* effector repertoire was specifically upregulated during biotrophic development and is one of the most highly expressed fungal genes during plant colonization, a good candidate for further investigation (Djamei et al., 2011). This enzyme was shown to be taken up by plant cells to redirect their metabolome in favour of the parasite. The authors proposed that the translocated fungal enzyme could act in conjunction with the maize chorismate mutase ZmCm2 in the plant cytosol by increasing the flow of chorismate from the plastid to the cytosol and in turn lowering the available substrate for salicylic acid biosynthesis in plastids. Salicylic acid enhances resistance of maize towards *U. maydis*, thus suppression of salicylic acid levels is likely to be a strategy to suppress plant defence for the success of the pathogen (Djamei et al., 2011).

High-throughput screens for *in planta* effector function are another approach to investigate roles for effectors and their contributions to pathogenicity. Suppression of plant defences is presumably one of the most important functions of effectors, so large-scale screens to detect such roles have been developed, especially for oomycete effector genes. For example, *Hyaloperonospora arabidopsis* RxLR effectors were delivered via the bacterial TTSS of the luciferase-expressing *Pseudomonas syringae* pv *tomato* DC3000 strain (Pst-LUX), and their ability to confer virulence assessed by measuring changes in Pst-LUX growth *in planta* with respect to controls (Fabro et al., 2011). The objective of this study was to test a number of effector candidates to investigate if they might play a role in suppressing PAMP-triggered immunity, or the second line of plant defence, effector-triggered immunity (discussed later). The results showed that the majority of the effector candidates contributed positively to Pst-LUX growth, while some decreased Pst-LUX growth on particular Arabidopsis accessions, and some conferred enhanced Pst-LUX growth on turnip (a non-host for *H. arabidopsidis*). This kind of information can help to rank the potency of effectors prior to more mechanistic investigations. Bioinformatic analyses of the draft genome sequences of *Phytophthora sojae*, *Phytophthora ramorum*, *Phytophthora infestans*, and *Hyaloperonospora arabidopsidis* (Tyler et al., 2006, Haas et al., 2009, Baxter et al., 2010) have identified a large superfamily of candidate effector proteins (134 to 565 per genome), with sequence similarity to the oomycete avirulence genes, *Avr1b*-1 (Shan et al., 2004), *Avr3a* (Armstrong et al., 2005), *ATR1* (Rehmany et al., 2005) and *ATR13* (Allen et al., 2004). A systematic functional characterization of a large sample set of the avirulence homolog genes encoded by the *P. sojae* genome was done by *Agrobacterium tumefaciens*–mediated transient expression in *Nicotiana benthamiana* (Wang et al., 2011b). The effectors were tested to
determine if they could suppress plant cell death triggered by different factors. Most of the effectors could suppress programmed cell death triggered by BAX (an animal programmed cell death-inducing protein), avirulence factors, and/or the elicitin INF1, while several effectors triggered cell death in the absence of other stimuli. Effectors were divided into subsets designated principal or secondary effectors according to their apparent contribution to the pathogen’s virulence. The expression of several principal effectors appeared to be timed to suppress first effector-triggered immunity and then PAMP-triggered immunity (discussed later) (Wang et al., 2011b). Another approach to improve the understanding of effector proteins consists of testing the ability of effector proteins to manipulate specific physiological processes or signaling pathways in the host cell. Very recently, *P. sojae* effectors were tested for the capacity to suppress RNA silencing in plants (Qiao et al., 2013). Individual effector and *GFP* genes were co-expressed by *Agrobacterium tumefaciens* infiltration in the leaves of *N. benthamiana* line 16c, which undergoes a systematic RNA silencing of *GFP* under the control of the cauliflower mosaic virus 35S promoter. In the absence of effector genes, both endogenous and exogenous *GFP* genes are silenced, resulting in no or very low green fluorescence in the infiltrated zone. Using this assay, two of 59 *P. sojae* RXLR effectors suppressed *GFP* silencing (Qiao et al., 2013). This study showed that oomycete pathogens have evolved effectors to facilitate infection by suppressing host RNA silencing. Further efforts to identify and characterize RNA silencing suppressors produced by other plant fungal pathogens will clarify if this is a common strategy used for plant pathogens to promote infection. Overall, identifying effectors with important and non-redundant virulence functions that constrain their evolution is important in that it helps to understand the diseases at the molecular level, and may allow deployment of more durable resistance sources to safeguard world food production (Dodds, 2010).

### 1.5. PLANT DEFENCE

In spite of the constant exposure of plants to a wide variety of potentially pathogenic microorganisms, disease is rare. Only well-adapted microbe species are able to infect target plant species and cause disease. Plants, unlike mammals, lack mobile defender cells and a somatic adaptive immune system, relying on the innate immunity of each cell and on systemic signals emanating from infection sites (Jones and Dangl, 2006). Plant innate immunity involves both preformed barriers and a multitude of responsive processes triggered by the recognition of potential pathogens, which leads to transcriptional reprogramming to respond to the attack. Individual plant species are resistant to the large majority of pathogens based on a general surveillance system specialized in detection of conserved microbial elicitors. These elicitors are typically essential
components of whole classes of pathogens, such as bacterial flagellin or fungal chitin, and are broadly called pathogen-associated molecular patterns (PAMPs) (Dodds and Rathjen, 2010). The plant surveillance system consists of receptor proteins called pattern recognition receptors (PRRs) that can be stimulated by PAMPs or alternatively by endogenous molecules released by pathogen invasion, such as cell wall or cuticular fragments, which leads to PAMP-triggered immunity (PTI) (Figure 1.4) (Dodds and Rathjen, 2010).

**Figure 1.4. Plant defence.** Molecules released from the pathogens into the extracellular spaces, such as the PAMPs flagellin and chitin, are recognized by cell surface pattern recognition receptors (PRRs) inducing PAMP-triggered immunity (PTI). Bacteria and fungi/oomycete pathogens deliver effector proteins into the host cell cytoplasm by specialised structures such as the TTSS and the haustorium respectively. These intracellular effectors often act to suppress PTI, but many can be recognized by intracellular resistance proteins (R), which induce effector-triggered immunity (ETI).

Successful pathogens are able to overcome PTI by means of secreted effectors that suppress PTI responses, becoming involved in a new battle where the detection of pathogen effectors elicits a second layer of plant defence (Figure 1.4) (Thomma et al., 2011). The second class of perception involves recognition by intracellular receptors of the pathogen’s effectors, inducing effector-triggered immunity (ETI) (Figure 1.4) (Dodds and Rathjen, 2010). The majority of these intracellular receptors are resistance proteins (R) of the nucleotide binding–leucine-rich repeat (NB-LRR) type that activate ETI. This mode of recognition leads to co-evolutionary dynamics between the plant and pathogen that are quite different from PTI as, in contrast to PAMPs, effectors are characteristically variable and dispensable (Dodds and Rathjen, 2010). As a result of selection pressure, pathogens evolve to lose or alter the effector that is recognized or to gain novel effectors to suppress the ETI response (Thomma et al., 2011). Plant receptors also evolve to recognize altered or newly acquired effectors, again resulting in ETI. This coevolution proceeds, with continuous selection for novel
pathogen isolates that overcome ETI and new plant genotypes that resurrect ETI (Thomma et al., 2011). The outcomes of PTI and ETI are generally similar, although ETI occurs quicker, is more prolonged and more robust than PTI, and often involves a form of localized cell death called the HR (Tsuda and Katagiri, 2010). PTI is generally effective against non-adapted pathogens in a phenomenon called non-host resistance, whereas ETI is active against adapted pathogens. However these relationships are not exclusive and depend on the elicitor molecules present in each infection.

1.5.1. Resistance against stripe rust in wheat

Resistance against stripe rust in wheat can be classified as race specific or race non-specific, and based on the plant growth stage it can be classified as seedling (also known as all-stage) or adult plant resistance (Chen, 2005). Seedling resistance is usually race specific, whereas adult plant resistance is usually non-race specific and often requires high temperatures to function. Race-specific resistance is often called qualitative as it relies on single major \( R \) genes (ETI). Major genes are effective in that they provide “resistance” rather than “tolerance” and their isolation and characterization is still a major goal in wheat research (Lowe et al., 2011). These genes are relatively easy to identify and deploy by breeding, but on the negative side they frequently lack durability (Lowe et al., 2011). Non-race specific resistance is also called quantitative in that it relies on genes with mainly minor individual effects on disease development, and these are often reported to have additive and/or epistatic effects (Vazquez et al., 2012). This resistance, characterised for being moderate or as slowing disease progress, is generally broad-spectrum (confers resistance to several unrelated pathogens) and durable. The environment can have a large influence on the effectiveness of non-race specific resistance, and the disease level on the same wheat variety may vary greatly in different environments (Vazquez et al., 2012). The effect of environment is particularly evident for a type of resistance designated ‘high temperature adult plant resistance’ (HTAP), where the resistance become more effective with increasing temperatures (Chen 2005). Approximately 53 officially named genes (\( Yr1-Yr53 \)), and many temporarily designated genes or quantitative trait loci (QTL) for stripe rust resistance, have been reported (Chen, 2005, Xu et al., 2013). Studies have shown that \( Yr11, Yr12, Yr13, Yr14, Yr16, Yr18, Yr29, Yr30, Yr34, Yr36, Yr39, Yr46, Yr48, Yr49, \) and \( Yr52 \) are adult-plant resistance genes (Bariana et al., 2006, Chen, 2005, Herrera-Foessel et al., 2011, Lowe et al., 2011, Ren et al., 2012), of which \( Yr18 \) and \( Yr36 \) have been cloned (Fu et al., 2009, Krattinger et al., 2009). However, since adult-plant resistance genes have only partial effects when present alone, the resistance conferred by them may be incomplete. Taking into account the advantages and
disadvantages of seedling and adult resistance, it is desirable to combine both types of resistance to achieve high-level, durable and sustainable control of stripe rust (Chen et al., 2013)

1.6. Sequencing technologies

DNA sequencing was first reported in 1968, when Wu and Kaiser (Kaiser and Wu, 1968, Wu and Kaiser, 1968) measured incorporation of radiolabeled nucleotides by *Escherichia coli* DNA polymerase in reactions that extended the 3’ termini of DNA fragments to fill in the complementary cohesive end sequences and reported a partial sequence. The Sanger and Maxam-Gilbert sequencing methods, which provided useful sequence analysis to standard laboratories, did not become widely used until ~1977 (Sanger et al., 1977, Maxam and Gilbert, 1977). The Sanger method is based on the selective incorporation of chain-terminating deoxynucleotides by DNA polymerase during in vitro DNA replication sequencing, and was the most widely DNA sequencing method for about 30 years. In the last decade, a number of different sequencing methods have been developed in order to overcome the limitations of Sanger sequencing with respect to throughput and costs (Service, 2006). Nowadays, the so-called second-generation or next-generation sequencing (NGS) technologies have not only revolutionized the field of genome sequencing but have also change the way we approach projects to understand how organisms function, develop, evolve, interact, and respond to environmental cues. In most NGS approaches, template DNA is fragmented, bound to a solid support, and amplified by PCR to generate clonal representations of the original fragments that are spatially separated for subsequent sequencing (Metzker, 2010). The sequencing itself is achieved by a number of methods that make use of different enzymes (polymerases or ligases) and chemistries to generate light signals that are recorded by highly sensitive detection methods (Nowrousian, 2010). What all methodologies have in common is that they generate sequence reads on a massive scale, without the requirement for DNA cloning, and at a fraction of the costs required for traditional sequencing (Marguerat et al., 2008). The current sequencing revolution is currently driven by three commercially available platforms: 454 (Roche), Genome Analyzer GAIIx and HiSeq 2000 (Illumina/Solexa) and ABI-SOLiD (Applied Biosystems) (Schuster, 2008). The first NGS technology that became commercially available was the Roche/454 genome sequencer (454 Life Sciences, Branford, CT) (Margulies et al., 2005). This method is based on the “sequencing by synthesis” principle or pyrosequencing. It differs from Sanger sequencing in that it relies on the detection of pyrophosphate release on nucleotide incorporation rather than chain termination with deoxynucleotides (Ansorge, 2009). The great advantage of this technology is the generation of relatively long single reads, currently offering read lengths of up to 1 Kb (www.454.com). This helps
with a variety of applications including de novo whole-genome sequencing, re-sequencing of whole genomes and target DNA regions, metagenomics, and RNA analysis (Ansorge, 2009). In genome sequencing projects, 454 can be useful to resolve larger structural rearrangements (insertions, deletions, inversions), as well as repetitive regions. The transcriptome data derived from 454 is ideally suited to detailed transcriptome investigation in the areas of novel gene discovery, gene space identification in novel genomes, assembly of full-length genes, single nucleotide polymorphism (SNP), insertion-deletion and splice-variant discovery (Ansorge, 2009). However it lacks read depth/coverage that are available with competing technologies.

Towards the end of 2006, the Illumina Genome Analyzer sequencing machine was introduced to the market, marking a change in genome sequencing in which short reads became the dominant technology for genomic applications (Fedurco et al., 2006). The most important feature was its massively parallel capacity for reading up to 80 million DNA templates simultaneously, even though at that time reads of only ~36–50 bp from each template were possible (Barski et al., 2007, Johnson et al., 2007, Mikkelsen et al., 2007). The technology is based on the principle of sequencing-by-synthesis chemistry that employs reversible chain terminators with removable fluorescent tags for each base (Ansorge, 2009). Unlike pyrosequencing, the DNA can only be extended one nucleotide at a time. Read length and throughput have undergone rapid improvement in the last few years, from 35 bp length reads with 1 Gb throughput using the Genome Analyzer, to protocols now available yielding 100 bp reads with 600 Gb output using the HiSeq 2000 instrument (Minoche et al., 2011). Illumina sequencing usually generates 50 times more reads than the 454 sequencing in a single run, at which point it becomes quantitative, making this technology ideal for applications requiring greater coverage such as ChIP-Seq, transcript counts, SNP detection, and small RNA analysis (Zhou et al., 2011). Each platform has its own advantages and limitations, but the differences in the chemistry of these two technologies makes each more suitable for different types of projects. Recent studies have showed that a combination of 454 and Illumina sequencing enables rapid and affordable genome sequencing (Seabury et al., 2013, Jiang et al., 2011, Dalloul et al., 2010) and de novo transcriptome studies for non-model species (Hornett and Wheat, 2012, Wenger and Galliot, 2013, Looso et al., 2013, Cahais et al., 2012)

1.6.1. *Pst* genomic resources

The stripe rust fungus lacks many biological characteristics that could make it a model system for molecular and genetic analyses. In *Pst*, unlike the stem rust and leaf rust fungi, the contribution of its
sexual stage to the evolution of pathogen virulence and aggressiveness is not yet well defined (Wellings, 2011). Pst is largely uncultivable *in vitro* and to date, no stable genetic transformation system for this organism has been established. At the time this research project started, there was a vast amount of literature available on stripe rust epidemiology, race dynamics, population structure and virulence evolution (Chen, 2005, Chen, 2007, Chen et al., 2002, Murray et al., 2005, Wellings, 2010b, Wellings and Kandel, 2004, Wellings and McIntosh, 1990), but very limited knowledge on the molecular strategies of this fungus for pathogenesis. In previous years, the main approaches undertaken to study the molecular genetics of Pst were through expressed sequence tags (EST) and full-length cDNA libraries. The first large scale transcript collection from Pst was reported by Ling et al. (2007). A full-length cDNA library of 42,240 clones was generated from uredospores of the Pst strain *PST*-78 (Ling et al., 2007). Among 196 cDNA clones selected randomly for DNA sequencing, only 73 of them have homologs of known genes and most of them are involved in housekeeping activities. Later, Zhang et al. (2008) tried to identify genes expressed during early infection stages, using EST analysis of germinated uredospores (Zhang et al., 2008). A total of 4,798 ESTs were sequenced from the germinated spores and assembled into 315 contigs and 803 singletons. Homology searches showed that the majority of these sequences had no similarity to known genes and did not shared homology with known fungal pathogenicity/virulence factors. More recently, Yin et al. (2009) constructed a cDNA library from Pst haustorial tissue to gain insight into haustorium-related functions. More than 5,000 ESTs from isolated Pst haustoria were generated and assembled into 1,134 unique sequences, from which only 10 showed homology to known proteins and 15 were predicted to be secreted from haustoria (Yin et al., 2009). Wang et al. (2009) studied Pst and wheat genes in a compatible interaction at different time points of the infection (ranging from uredospore germination to the generation of sporogenous cells) by the use of cDNA-amplified fragment length polymorphism (cDNA-AFLP). This method allowed the identification of 186 genes (9 from Pst) differentially expressed in different stages of the infection process (Wang et al., 2009b). Very interestingly, in parallel to this research, other studies were pursued on Pst to generate a draft genome sequence (Cantu et al., 2011) and to uncover the potential effector gene candidates encoded in its genome (Cantu et al., 2013). As part of this study, I generated and analyzed a vast amount of sequencing data generated by NGS to uncover effector gene candidates expressed in the haustorium and gain a better understanding of the molecular strategies of Pst in disease causation (Chapters 2 and 3, (Garnica et al., 2013)). Overall, an unprecedented amount of genomic resources were generated for Pst in the last three years, which has enormous potential to explore the pathogenic mechanisms of this important agricultural wheat pathogen.
1.7. PROJECT AIMS AND EXPERIMENTAL APPROACH

Understanding the strategies by which rust fungi establish compatible interactions with their hosts requires uncovering the molecular events accompanying the key pathogenic stages. The objective of this study was to investigate the molecular strategies of *Puccinia striiformis* f.sp. *tritici* by deep transcriptomic and proteomics analysis of the early infective and haustorial stages of the pathogen, with special emphasis on the prediction and characterization of effector genes. In the absence of any existing genomic or transcriptomic data for *Pst*, the transcriptomes of isolated haustoria and germinated spores were generated by 454 RNA-seq, effector gene candidates were predicted from the assembled sequences, analyzed for homology to known genes, expression profiles, and sequence content, and methods for their functional characterization were established (Chapter 2). Illumina RNA-seq data was generated from isolated haustoria and germinated spores to allow comparative transcriptome analysis of the two pathogenic stages, and through extensive annotation of the transcripts, marked metabolic differences between them were established (Chapter 3). A new technique to isolate highly purified haustoria from infected tissue was developed by combining Percoll density gradients and fluorescence-activated cell sorting (Chapter 4), which was used to generate proteomic data from this important pathogenic state (Chapter 5). To validate the transcriptomic studies and effector prediction, as well as to assist the annotation of a high quality *Pst* assembled genome (from a parallel project in this laboratory), additional proteomic data from ungerminated and germinated spores was generated, and preliminary analyses of the data were carried out (Chapter 5).

Through the success of this project, we now have a detailed picture of the strategic metabolic switch of *Pst* during the early to the advanced stages of the infection, in addition to a complete catalog of annotated transcripts (many with proteomic data support), which constitute a unique resource for future research. Apart from the biological insights, this project also required the development of new methods to study a non-model, biotrophic plant pathogen. As a result, we now have the ability to isolate ultrapure *Pst* haustoria, which was not possible before this study. Finally, this project pioneered the search for *Pst* effector candidates expressed in the haustoria, generating a comprehensive set of candidate effector genes whose future characterization will improve the understanding of *Pst* pathogenic strategies.
Chapter 2

Prediction and characterization of \textit{Pst} effectors

2.1. INTRODUCTION

Haustoria are believed to be the primary site of nutrient uptake from the host (Hahn et al., 1997a, Voegele and Mendgen, 2003), and have also been associated with the secretion of virulence proteins called effectors (Panstruga and Dodds, 2009). Effector proteins from biotrophic fungi are important virulence factors deployed by the pathogen during its interaction with plant host tissue (Stergiopoulos and de Wit, 2009). They travel through the canonical eukaryotic endomembrane secretion pathway for deposition outside of the fungal cell (Ellis et al., 2009, O'Connell and Panstruga, 2006). Very little is known about the effector activities of eukaryotic pathogens, but broadly they are thought to manipulate the physiological and immune responses of host cells during infection. Conversely, in certain cases the host has evolved to recognize effectors through the action of specific intracellular immune receptors. In this case, the recognized effector is called an Avirulence protein (Avr), and serves as a signal for the plant to induce defences to block pathogen growth (Gohre and Robatzek, 2008).
Chapter 2: Prediction and characterization of Pst effectors

The identification and characterization of effectors secreted by pathogens during infection contributes to our understanding of strategies for pathogenicity. There are no formal criteria to define effector genes, but effectors are loosely defined as molecules with the capacity to alter the structure and function of the host cell during infection (Hogenhout et al., 2009). Effectors are typically highly variable in sequence and the majority do not have known homologs, which makes their identification through sequence analysis difficult (Duplessis et al., 2011, Saunders et al., 2012). The recognition of Avr proteins by the plant immune system has been one successful approach for map-based identification and cloning of effector genes (Orbach et al., 2000, Dodds et al., 2004, Ridout et al., 2006). Effector proteins typically carry an N-terminal secretion signal, a characteristic that has been exploited for effector gene identification in pathogen EST libraries (Hahn and Mendgen, 1997, Catanzariti et al., 2006, Mosquera et al., 2009). Other characteristics as small size, high cysteine content and conserved sequence motifs have been proposed as additional criteria for their identification (Stergiopoulos and de Wit, 2009), although none of these criteria are essential (Martin and Kamoun, 2012). With the introduction of high-throughput sequencing technologies, the sequencing of whole genomes has become routine, which combined with computational pipelines enables the in silico prediction of effector genes (Mueller et al., 2008, Duplessis et al., 2011). Recent studies have also used high-throughput transcriptomic data from plant pathogenic fungi for large scale prediction of effector genes ((Wessling et al., 2012, Fernandez et al., 2012), N. Upadhyaya, CSIRO Plant Industry, Personal communication), resulting in datasets that not only list predicted genes but also contain information about their expression.

Wheat rust pathogens, including Pst, Pgt and Pt, are a very important group of biotrophic fungal pathogens that pose a major threat to wheat production worldwide (Dean et al., 2012). Pst and Pgt have huge worldwide economic impacts every year (Dean et al., 2012). Thus, the study of these diseases is fundamental to the development of control methods. Considerable efforts have been made to generate a high quality genome assembly for Pgt, from which a large number of effector genes have been predicted ((Duplessis et al., 2011) and N. Upadhyaya, CSIRO Plant industry, personal communication). However, despite the agricultural importance of Pst, molecular studies on this pathogen were scarce at the time I started this project and its effector genes were largely uncharacterised.

During the course of this research, a draft genome of an American Pst strain was published, consisting of ~30,000 contigs (DNA sequence assemblies representing overlapping segments derived from a sequencing project) with very little genetic annotation (Cantu et al., 2011). At the completion
of the work described in this Chapter, two further draft _Pst_ genomes and a list of predicted effector candidates from two British _Pst_ strains were also published (Cantu et al., 2013). The isolate chosen for this project was _Pst_ 104 E137 A-, the first incursion of wheat stripe rust detected in Australia (O’Brien et al., 1980). Over the last 30 years, 19 new pathotypes derived from this founder strain have arisen in Australia, apparently as a result of single gene mutations (Wellings, 2007). In this scenario, it is advantageous to work with _Pst_ 104 E137 A- as the information derived from this work can be used for future studies on the pathogen’s molecular evolution.

In this chapter, haustoria were investigated as a source of potentially secreted proteins to identify candidate effector genes. This involved using a combination of haustoria isolated by affinity chromatography and next-generation sequencing (NGS) (Roche 454 platform) to generate the first _Pst_ haustorial transcriptome, and from this effector genes were predicted. The use of different software to assemble the transcriptome and predict the putatively secreted proteins, followed by extensive manual curation, resulted in the first set of potential effector candidates of _Pst_ comprising 437 genes. The candidate effector genes were analyzed for homology to known genes, and the presence of motifs or structural features in their nucleotide and predicted protein sequences. Additionally, I developed and applied functional tests based on recognition of effector candidates by wheat cultivars carrying different resistance genes, and the capacity of individual effectors to inhibit plant cell death, to a subset of the candidate genes. I also generated transcriptomic data from germinated spores. Thus, differential expression of effector candidates using germinated spores as control tissue was possible. The large amount of data generated on the Illumina platform allowed rigorous statistical measures of gene expression, and the classification of effector candidates into four expression groups. This classification, together with the other aspects studied in this Chapter, provides valuable information to prioritize effectors for further functional characterization, as well as a better understanding of the molecular strategies of _Pst_ in deploying its effectors to cause disease.

2.2. MATERIALS AND METHODS

2.2.1. Fungal and plant material

_Pst_ strain 104 E137 A- was provided by Dr. Colin Wellings from the Plant Breeding Institute, University of Sydney, Cobbity NSW, as uredospores. For multiplication of spores and for haustoria isolation, seedlings of the _Pst_-susceptible wheat ( _Triticum aestivum_ ) cultivar (cv.) Morocco were
grown in growth chambers at 21°C and 16:8 h light:dark cycle. Plant cell death inhibition assays on wheat were performed using wheat cv. Grandin. Avr recognition assays were done using wheat cvs. Chinese 166, Lee, Heines Kolben, Moro, Clement, Heines Peko, Compair, Carsten V and Spalding prolific, all provided by Dr. Colin Wellings.

Plants were grown in a mixture of soil, pine bark and sand (3:2:1) and the plant growth regulator maleic hydrazide (Agency, 1994), was applied to wheat seedlings (~3 cm tall) five days after sowing, 15 ml per pot (10 cm diameter x 15 cm high) at a concentration of 1.1 g/L, unless otherwise indicated. Plant cell death inhibition assays infiltration controls were performed on *Nicotiana tabacum* (tobacco) transgenic plants expressing the flax rust resistant gene *M* (Catanzariti et al., 2006).

### 2.2.2. Infection of wheat with *Pst*

#### 2.2.2.1. Strain purification by single pustule isolation

Uredospores collected from the first stripe rust incursion in Australia were donated by Dr Colin Wellings and Prof. Robert Park (University of Sydney, Cobbitty); this strain is designated 104 E137 A- (Steele et al., 2001). To propagate this strain clonally, spores were activated by a heat-shock treatment of 3.5 min at 42°C. Approximately 10 mg of heat shock-treated spores were mixed with talcum powder (1:7 w/w) and were sprayed homogenously with a manual air pump onto seven day old wheat seedlings previously wetted with water. Plants were maintained in wet chambers (100% relative humidity) at 9°C for 24 h in darkness, then transferred to a constant-temperature growth cabinet at 17°C with a 16:8 h light cycle. Once leaves became symptomatic (usually 9 days after infection, 9 dai), leaves containing visible single infection spots separated by at least 6 cm were detached and processed immediately. Superficial dirt and talc particles were wiped off and working under sterile conditions, detached leaves were surface-sterilized as shown in Figure 2.1. Briefly, leaves were submerged sequentially in sterile Milli-Q (MQ) water for 1 min, in 2% (v/v) sodium hypochlorite for 3 min, rinsed in sterile MQ water for 2 min, then rinsed in 70% ethanol for 1 min and sterile MQ water for 2 min. After the final washing step, remaining moisture was wiped off and leaves were cut into pieces (~5 cm) containing a single infection spot. Individual leaf pieces were anchored upright in solid Murashige and Skoog media (MS) (Sigma, St. Louis, MO) supplemented with 1% sucrose, contained in sterile propylene histology jars, and were maintained at 17°C with a 16:8 h light cycle for 3 days or until pustules were visible. The jars were carefully laid down horizontally and sterile aluminum foil pieces of 3 cm x 3 cm were placed at the bottom of the jar so
that the leaves were suspended above it, allowing newly produced spores to fall onto the foil. Spores from a single pustule were collected at least three times every 2-3 days, for as long as the MS media remained sterile. Clonal spores obtained from each single pustule were multiplied by sequential infections of cv. Morocco wheat plants. For long term storage of collected spores, 200 mg aliquots were vacuum-dried for 30 min at room temperature and then stored at -80°C.

**Figure 2.1. Single pustule isolation.** 1. A symptomatic leaf containing spaced single pustules is bleach-ethanol sterilized. 2. Leaf tissue is cut into pieces each containing a single pustule. 3. Each tissue piece is anchored in solid MS media and kept at 17°C until sporulation.

### 2.2.2.2. Plant infection and collection of infected tissue

Approximately 150-200 mg of dormant spores were heat-shock activated as in Section 2.2.2.1 unless freshly collected spores were available, in which case the activation treatment was omitted. Batches of approximately 360 seven-day old wheat seedlings (~15 seedlings/pot) were sprayed with MQ water to wet the leaf surfaces, and subsequently a mixture of spores and talcum powder (1:5 w/w) was sprayed homogenously onto the seedlings. Plants were maintained in wet chambers (100% relative humidity) at 9°C for 48 h in darkness, although leaves were sprayed with MQ water after 24 h to increase the extent of infection. Infected seedlings were moved to a constant-temperature growth cabinet at 17°C with a 16:8 h light cycle. Nine days after infection, only those areas with visible infection spots were collected and processed immediately for haustoria isolation.

### 2.2.3. Haustoria isolation and transcriptome sequencing (RNA-seq)

#### 2.2.3.1. Haustoria isolation by Concanavalin A (ConA) affinity chromatography
2.2.3.1.1. Coupling of CNBr-activated Sepharose 6MB to ConA

CNBr-activated Sepharose 6MB (4 g; GE Healthcare) was hydrated in ~800 ml of 1 mM HCl for 15 min. The beads were washed with 60 ml of coupling buffer (0.1 M NaHCO$_3$, 0.5 M NaCl, pH 8.3) then immediately added to 60 mg of ConA (Amersham Biosciences) dissolved in 14 ml of coupling buffer, then mixed by inversion on a rotary wheel for 2 h at room temperature or overnight at 4°C. After incubation, the beads were washed extensively with ~240 ml of coupling buffer to remove excess ConA. The unconjugated active groups on the beads were blocked by adding 40 ml coupling buffer containing 0.2 M glycine, and mixed for a further 2 h at room temperature on a rotary wheel. The ConA-Sepharose beads were washed three times alternating with 0.1 M Sodium Acetate, 0.5 M NaCl pH 4.0 followed by 0.5 M NaCl, 0.1 M Tris-HCl pH 8.0 (150 ml each wash). Finally the beads were equilibrated with 200 ml of storage buffer (0.15 M NaCl, 10 mM Tris-HCl pH 7.2, 1 mM CaCl$_2$, 1 mM MnCl$_2$, 0.02% (w/v) NaN$_3$) and stored at 4°C in the same buffer, so that the volume above the beads was equal to the beads bed volume. All washes were carried out in a funnel with a sintered glass filter under vacuum filtration.

2.2.3.1.2. Regeneration of column after use

Between each use, the beads were recovered from the column and washed on a sintered glass filter three times, alternating between 0.1 M Sodium Acetate, 0.5 M NaCl pH 4.0 and 0.5 M NaCl, 0.1 M Tris-HCl pH 8.0 (10 ml each per ml of beads). The beads were then equilibrated with storage buffer and stored at 4°C.

2.2.3.1.3. Isolation of haustoria from Pst-infected wheat leaves

Heavily infected wheat leaves (~20-25 g) were harvested 9 dai (leaf pieces of ~6 cm length). To remove superficial traces of spores, talc and contaminating organisms, leaves were washed extensively with tap water, then incubated for 3 min in 2% bleach (v/v). Bleach was removed by washing with tap water three times, and the leaves were incubated for one min in chilled 70% ethanol. Ethanol was removed by washing with chilled MQ water several times and the infected tissue was dried as much as possible with sterile paper towel. All subsequent steps were carried out at 4°C or on ice. Infected plant material was homogenized using a Waring blender in 180 ml of cold homogenization buffer (0.3 M sorbitol, 20 mM MOPS pH 7.2, 0.1% (w/v) BSA, 0.2% (w/v) PEG 6000,
0.2% (v/v) β mercaptoethanol (added freshly)) at maximum speed for 25 seconds. The homogenate was filtered first through a 100 µm nylon mesh, and then through a 20 µm mesh, by gravity flow. The filtrate was distributed in six chilled Oak Ridge round-bottom centrifuge tubes, 45 ml capacity, and centrifuged at 6,900 \( g \) for 5 min using a refrigerated benchtop centrifuge with a swingout rotor. Pellets were resuspended in 6ml of suspension buffer (0.3 M sorbitol, 10 mM MOPS pH 7.2, 0.2% (w/v) BSA, 1 mM KCl, 1 mM MgCl₂, 1 mM CaCl₂) and kept on ice. ConA-Sepharose beads were loaded onto three Glass Econo-Columns (1.5 x 10 cm, 18 ml; Bio-Rad), 5 ml bed volume per column, and were equilibrated with suspension buffer. Each column was then loaded with 1 ml of suspended pellets, allowing the suspension to migrate into the beads to be incubated for 15 min with stopped flow. Columns were gently rinsed with 5 ml of suspension buffer and the incubation was repeated with the remaining suspension. Each column was washed by carefully layering 10 ml of suspension buffer on top of the ConA-Sepharose beads and then allowing the buffer to flow through until it ran clear. Immobilized haustoria were released by adding 3 ml of suspension buffer and vigorously agitating the beads by pipetting up and down using a 1 ml pipette tip with the end cut off. Immediately after the beads settled, the supernatant containing the released haustoria above the beads was collected, and this step was repeated three times per column. All supernatants were combined (keeping a fraction for microscopic analysis), centrifuged at 14,700 \( g \) for 5 min and after discarding the supernatant, the pellet was snap frozen in liquid nitrogen and stored at -80°C.

### 2.2.3.2. Total RNA isolation from purified haustoria and germinated spores

Frozen haustoria pellets were ground with help of an electrical tissue homogenizer, and total RNA was isolated using the QIAGEN (Doncaster Australia) Plant RNeasy kit following the manufacturer’s instructions, processing three pellets per column. Each column was eluted in 53 µl of RNase-free water and samples were checked for concentration and purity using a NanoDrop® ND-1000 UV-Vis Spectrophotometer. RNA samples were combined in pairs and treated with DNase I (New England Biolabs, Beverly, MA) following the manufacturer’s instructions. All samples were purified from DNA and proteins using the QIAGEN (Doncaster Australia) Plant RNeasy kit following the cleanup protocol. Each column was eluted in 50 µl of RNase-free water and checked for concentration on a NanoDrop ND-1000 UV-Vis Spectrophotometer.

To extract RNA from spores, ~100 mg of freshly collected spores were germinated by evenly spraying them onto MQ water contained in a large glass petri dish (15 mm diameter, 8 cm high) and incubating them in darkness for 15 h at 9°C. Germination was verified by microscopic examination, and germinated spores were passed through a 15 µm nylon mesh to remove water. Retained
material was dried as much as possible with tissue paper and snap frozen in liquid nitrogen. Three samples were ground in liquid nitrogen, and total RNA isolation, DNAse I treatment and RNA clean-up were performed as for haustoria tissue.

2.2.3.3. mRNA isolation and pyrosequencing

Messenger RNA was isolated from total RNA using Dynabeads® mRNA Purification Kit (Invitrogen, Cat No. 610.06) following the manufacturer’s instructions. For haustoria, all samples of total RNA were combined and the kit reagents were increased proportionally. mRNA was eluted in 20 µl RNase-free water and concentration and integrity were verified on an Agilent 2100 bioanalyzer and also by using the Ribogreen quantitation assay (Promega). mRNA from germinated spores was processed in the same way. Haustoria and germinated spores mRNA samples were sent to the Biomolecular Resource Facility (Australian National University, Acton, Australia) where they were fragmented into a 200-1500 nt population with a 500 nt peak which was subsequently purified. The fragmented population was used as template and primed with random hexamer primers (primer “random”, Roche Applied Science) to synthesize double stranded cDNA. Subsequently, adaptors were blud-end ligated to cDNA, the library was size selected and pyrosequencing (single-end strategy) was carried out on a Roche 454 Titanium sequencer (454). One half of a 454 plate was used for each sample following standard procedures recommended by Roche. Initial quality filtering of the raw reads was performed using the Roche proprietary analysis software Newbler (software release 2.0.00.22) to remove poor quality reads and adapter sequences.

2.2.4. De novo assembly of haustoria and germinated spores transcriptomes, and prediction of effector gene candidates

CLC Genomics Workbench 4.0 software (http://www.clcbio.com/) was used for de novo assembly of 454 reads and prediction of open reading frames (ORFs). The parameters used for de novo assembly were similarity 0.97; length fraction 0.5; insertion cost 3; deletion cost 3; mismatch cost 2. The parameters to predict ORFs from the clean haustorial transcriptome assembly were: starting codon ATG, both strands searched, allow open end sequence (consider ORF even if stop codon is not contained in the transcript) and minimum length 80 amino acids (aa). Predicted ORFs were translated into proteins and analyzed further with SignalP 3.0 and 4.0 (BendtSEN et al., 2004) for detection of signal peptides for protein secretion. The resulting list was analyzed with TMHMM2.0 (Krogh et al., 2001) to discard those predicted secreted proteins containing transmembrane domains.
which would represent integral membrane proteins. SignalP and TMHMM were both run with default parameters. The set of predicted putative effector candidates was additionally analyzed with TargetP software (Emanuelsson et al., 2007) to predict protein subcellular localization, and proteins with mitochondrial targeting signals were discarded.

2.2.5. Characterization of candidate effector genes

2.2.5.1. Validation of expression of candidate effector genes by reverse transcriptase-PCR (RT-PCR)

Total RNA was isolated from Pst-infected tissue (9 dai), uninfected wheat leaves, germinated spores, and isolated haustoria. All tissues were ground in liquid nitrogen and extracted using the QIAGEN (Doncaster Australia) Plant RNeasy kit according to the manufacturer’s instructions. For cDNA preparation, 2 µg of total RNA were mixed with 1 µl of oligo(dT)$_{18}$ and MQ water up to 11 µl and heated at 70°C for 10 min, followed by cooling on ice for one minute. DTT was added to the RNA to a final concentration of 1 mM together with 1 µl of dNTPs (each at 10 mM), 4 µl of 5X superscript II buffer (supplied with Superscript kit) and 0.5 µl of SuperScript II reverse transcriptase (200 units/µl, Invitrogen). After incubation for 1 h at 42°C, the reaction was stopped by heating for 15 min at 70°C, followed by the addition of 1 µl of RNase H (1 unit/µl) to each tube and incubated for 20 min at 37°C. Target cDNAs were PCR amplified using a dilution of 1:15 of the synthesized cDNA as template, specific forward and reverse primers (list of primers in Appendix 1), and 2x PCR Master Mix (Promega). PCRs were performed using a MiniCycler (MJ Research) and consisted of 30 cycles of denaturation and annealing of 58°C.

2.2.5.2 Annotation of candidate effector genes

The final set of candidate effector genes was annotated by searching public databases with the nucleotide sequence truncated to remove the portion encoding the predicted signal peptide. BLASTn and BLASTx searches against the NCBI nucleotide and non-redundant proteins databases were performed to find similarity with known genes/proteins in other organisms. Gene ontology (GO) classification was performed using BLAST2GO 2.5.1 (B2G) (Conesa et al., 2005), which categorized query sequences into classes including cellular components, molecular function, and biological processes, where possible. The parameters in B2G were set to maximum e-value < 10$^{-25}$, maximum number of alignments to report = 20 and highest scoring pair length = 33 aa. Motif content analysis of the translated putative effector set was done using the online interface of the MEME Suite (motif-
based sequence analysis tools, http://meme.nbcr.net/meme). The parameters used were optimum width of each motif within 3-50 aa, maximum number of motifs to find 100, and one or more occurrences of a single motif per sequence. Cysteine content and distribution analysis was done using CLC genomics and Microsoft Excel.

2.2.5.3. Cloning of candidate effector genes into a specialized vector (pEtHAn) for secretion via the type-III secretion system of Pseudomonas fluorescens

2.2.5.3.1 Cloning of candidate effector genes into the Gateway entry vector pENTR/D-TOPO

Primers to amplify the selected candidate effector genes were designed to amplify the full ORF excluding the sequence encoding the predicted signal peptide, and to have annealing temperatures equal to or slightly above 58°C (the same primers were used for RT-PCR assays, Appendix 1). The sequence CACCATG was included on the 5’ end of all the forward primers. The CACC sequence is necessary for directional cloning into the pENTR/D-TOPO vector (Invitrogen), and the ATG initiation codon was included to make the construct useful for future C-terminal fusion-tagged proteins. The reverse primers were designed based on the C-terminus of the ORF excluding the stop codon.

Complementary DNA from isolated haustoria was diluted 1:5 with water and was used as template to amplify the effector candidates by PCR. Phusion High-Fidelity DNA Polymerase (New England Biolabs) was used according to the manufacturer’s instructions, using 6.5 µl of template per 50 µl total reaction volume. PCRs consisted of 30 cycles of denaturation and annealing at 58°C for 30 sec. PCRs that produced a single band of the expected size were purified with the GenElute PCR Clean-Up Kit (Sigma). PCRs that produced a clear strong band of the expected size but also some weaker non-specific bands were resolved on a 1% agarose gel and the band of interest was cut out and purified with the GenElute PCR Clean-Up Kit. Purified PCR products were cloned into the pENTR/D-TOPO vector using the pENTR/D-TOPO cloning kit (Invitrogen) according to the manufacturer’s instructions, and transformed into 12.5 µl Shot TOP10 Chemically Competent E. coli cells (Invitrogen, Carlsbad CA). Colonies were screened by colony PCR using the gene-specific forward primer and a modified M13-reverse primer (TOPO #282: 5’-AGGAAACAGCTATGACCATGTAATACGACTCAC-3’). Miniprep cultures of selected positive colonies were processed with the Wizard Plus Minipreps DNA Purification System (Promega) and digested with restriction enzymes NotI and EcoRV (New England Biolabs), to confirm their identity. Confirmed positive clones were sequenced with M13-forward and modified M13-reverse and used as Gateway entry clones.
2.2.5.3.2. Recombination of candidate effector genes into the destination expression vector pNR526-C2AG3A

Expression clones were constructed with the Gateway cloning system LR Clonase® II enzyme mix (Invitrogen) according to the manufacturer’s instructions. The destination vector pEDV-AvrRPM1-G2AC3A-Des1 which contains a pBBR-1 origin of replication, the mobilization factor mob, a gentamicin resistance gene and the Gateway recombination site downstream of the bacterial type-III effector gene AvrRPM1 promoter and the 5’ portion of this gene encoding the type three secretion signal peptide, was a kind gift of Dr Narayana Upadhyaya (CSIRO Plant Industry, Black Mountain, ACT, Australia). The Gateway LR reactions (recombination reaction between attL and attR sites) were transformed individually into 12.5 µl aliquots of Shot TOP10 Chemically Competent E. coli cells (Invitrogen, Carlsbad CA), and bacterial colonies were screened by colony PCR with the gene-specific forward primer and a reverse primer based on pEDV-AvrRPM1-G2AC3A-Des1 downstream of the end of the HSP sequence (DMscreen2-R: 5’- TCACGACGTTGTAAAACGACGGCCAG-3’). Miniprep cultures of selected positive colonies were processed with the Wizard Plus Minipreps DNA Purification System (Promega) and the purified plasmids were digested with restriction endonuclease EcoRV (New England Biolabs), to confirm their identity. Confirmed clones were sequenced with a forward primer designed on pEDV-AvrRPM1-G2AC3A-Des1 upstream of the cloning site (DMscreen2-F: 5’- TCACGACGTTGTAAAACGACGGCCAG-3’) and DMscreen2-R.

2.2.5.3.3. Transformation of Pseudomonas fluorescens Pf0-1 (pEtHAn) bacteria by triparental mating

Effector candidates cloned in the expression vector pEDV-AvrRPM1-G2AC3A-Des1 were transformed into Pseudomonas fluorescens strain Pf0-1 (pEtHAn) (Thomas et al., 2009) by triparental mating. The strain P. fluorescens Pf0-1 pEtHAn was the receptor strain, E. coli HBA101 carrying the plasmid pRK2013 the helper strain, and E. coli TOP10 carrying each candidate effector gene in pEDV-AvrRPM1-G2AC3A-Des1 was the donor strain. The three strains were grown for ~16 h with agitation (250 rpm) in 2 ml LB broth containing the appropriate antibiotic (receptor strain at 28°C, 30 mg/ml chloramphenicol; helper strain at 37°C, 50 mg/ml kanamycin; donor strain at 37°C, 25 mg/ml gentamicin). All cultures were diluted with LB broth without antibiotics (receptor strain 1:5, donor and helper 1:100) and were incubated for 6 h with agitation (250 rpm) at the strain-appropriate temperatures to an OD_{600} of 1.5-1.8. Forty µl of each culture was mixed in a sterile eppendorf and 30
µl of the mixture was spotted on LB solid media without antibiotics and incubated at 28°C for 24 h. The resulting culture was sampled and streaked onto LB solid media containing 30 mg/ml chloramphenicol and 25 mg/ml gentamycin to select against the parent strains and to obtain single 
P. fluorescens Pf0-1 colonies containing the recombinant plasmid. After 48 h incubation at 28°C, a single colony was picked, re-streaked onto LB solid media containing 30 mg/ml chloramphenicol and 25 mg/ml gentamycin to obtain single colonies. A few single colonies for each construct were analyzed by colony PCR with the appropriate gene-specific forward primer and DMscreen2-R to verify the presence of the candidate gene, and also with primers specifically designed to amplify the 16S ribosomal subunit of 
P. fluorescens (forward 5’-CGAGCGAACGGGGACTAGCCCTTAAG-3’, reverse 5’-GACTAGACGCCCTATTAAGACTCGC-3’) to confirm the bacterial identity.

2.2.5.3.4. Activation of the 
P. fluorescens type-III secretion system for infiltration assays

Bacteria of 
P. fluorescens Pf0-1 (pEtHAn) carrying candidate effector genes were grown overnight (O/N) in 3 ml LB broth containing 30 mg/ml chloramphenicol and 25 mg/ml gentamycin at 28°C. The O/N cultures were used to inoculate 30 ml of LB broth containing 15 mg/ml gentamycin and were incubated at 28°C for 4-5 h to an OD$_{600}$ of 0.7-0.8. Cultures were centrifuged at 4000 g at 4°C for 6 min, and pellets were washed twice with 25 ml cold 10 mM MgSO$_4$. Pellets were resuspended in minimal media (50 mM potassium phosphate buffer, 7.6 mM (NH$_4$)$_2$SO$_4$, 1.7 mM MgCl$_2$, 1.7 mM NaCl, 10 mM fructose, pH 5.7) containing 15 mg/ml gentamycin to an OD$_{600}$ of 0.8 and were incubated overnight at 20°C. Cultures were centrifuged at 4000 g at 4°C for 6 min and pellets were resuspended in 10 mM MgCl$_2$ to an OD$_{600}$ of 1.0. Bacterial suspensions were used immediately for infiltration.

2.2.5.4. Effector recognition assay

Seeds from nine different wheat cultivars (see Section 2.2.1) were vernalised for 72 h at 4°C and pre-germinated in petri dishes lined with wetted filter paper. Pre-germinated seeds were transferred to soil and grown for 5 days in a 21°C growth chamber in a 16:8 h light:dark cycle. Seedlings were moved into a greenhouse room at 21°C with high relative humidity, without artificial light, and were allowed to acclimatize for 5 days before infiltration. Activated 
P. fluorescens Pf0-1 (pEtHAn) carrying candidate effector genes were syringe infiltrated into the primary and secondary leaves, and presence or absence of the hypersensitive response was scored 3 and 8 days after infiltration.
2.2.5.5. Plant cell death inhibition assay

A SnToxA clone provided by Dr Peter Solomon (RSB, ANU) was used to produce the necrosis toxin *Stagonospora nodorum* ToxA from *E. coli*, as described by Manning et al. (2004). Briefly, the recombinant N-terminally His-tagged ToxA (His-ToxA) was expressed heterologously in *E. coli* (clone donated by Dr Peter Solomon, RSB, ANU) by induction with IPTG for 20 h. Cells were lysed with lysozyme and repetitive freezing-thawing, and the recombinant protein was captured by incubating the cell extract with Ni-NTA resin (Qiagen). Bound proteins were eluted by sequential washes with 8 M urea buffers of pH 8.0, 6.3, 5.9 or 4.5, and all the fractions were collected. Samples from all fractions were run on a 16% SDS-PAGE gel to see in which the ToxA protein band of 17-18 kDa was present, and those fractions were pooled (Manning et al., 2004). The purified protein was refolded using urea step dialysis as described previously (Tuori et al., 2000). Protein concentrations were determined with the Detergent Compatible protein assay kit (Bio-Rad) using BSA as the standard. The ToxA protein obtained using this method had a concentration of 3.22 mg/ml. Activity tests were done on ten day old seedlings of the susceptible wheat cv. Grandin, which contains the susceptibility gene *Tsn1* (Liu et al., 2006). These plants were grown as in 2.2.1 and their leaves were infiltrated with serial dilutions of the purified toxin. The dilution that induced visible chlorosis 24 h after infiltration and tissue necrosis 72 h after infiltration was chosen for the rest of the assays.

To test the capacity of effector candidate genes to inhibit plant cell death triggered by ToxA, activated *P. fluorescens* Pf0-1 (EtHAn) carrying each effector candidate were syringe infiltrated into secondary leaves of nine day old seedlings of the wheat cv. Grandin. The diluted ToxA preparation was infiltrated 16 h later in the same leaf sections where effector candidates were previously infiltrated, and signs of necrosis were scored 3 days after ToxA infiltration.

2.3. RESULTS

2.3.1. *Pst* isolate purification

In contrast to other *Puccinia* species, *Pst* requires low temperatures for infection. Thus the first step was to establish infection procedures. A household refrigerator was adapted to maintain a constant temperature of 9°C and was used as incubator to induce rust infections. Although plants were initially incubated at 9°C for 24 h to allow spore germination and penetration, the amount of infected tissue was increased about 30% by extending this incubation period to 48 h and spraying the plants with water 24 h later (Figure 2.2). Infected plants were kept in a plant growth cabinet
maintained at 17°C with a 16:8 h light:dark cycle for the rest of the infection period. Under these conditions, the first symptoms (slightly white spotted leaves) were visible at 6 dai, and first pustule eruptions were visible 10 dai. Wheat cv. Morocco which is highly susceptible to most Pst Australian pathotypes, including Pst 104 E137 A- (Figure 2.2), was used for haustoria isolation and amplification of spores.

![Figure 2.2. Morocco wheat seedlings infected with Pst 104 E137 A-. Infected plants 17 dai maintained in a growth cabinet at 17°C, ready for collection of spores.](image)

Once the infection conditions were established, spores were purified using two different procedures. Plants were inoculated with a low amount of inoculum and observed closely after 7 dai to find an isolated non-erupted pustule on a section of a leaf that could support good spore development. All other leaves from the same pot were removed, and the pot was isolated by inserting it in a clear transpiration bag to allow air and moisture exchange while preventing contamination by external Pst spores. Although this procedure was attempted several times, the leaf carrying the pustule always became very chlorotic and the erupted pustule did not produce abundant spores. Thus, the isolation protocol was changed, and leaves carrying single pustules were detached from the plant 9 dai, surface-sterilized, anchored in MS media, and kept in sterile plastic jars at 17°C until pustule eruption. Spores from the same pustule were collected multiple times, producing enough inoculum for a second round of single pustule isolation which was done following the same method. The spores obtained by this method were then amplified to produce enough material for infecting plants on a regular basis.
2.3.2 Isolation of haustoria

Haustoria were isolated from infected wheat leaves by affinity chromatography based on the method described by Hahn and Mendgen (1992). This method utilizes Concanavalin A, a jack bean-derived lectin which has affinity for internal and non-reducing terminal α-D-mannosyl and α-D-glucosyl groups present in various sugars, glycoproteins, and glycolipids (Goldstein and Poretz, 1986). This lectin, linked to Sepharose beads, is thought to capture the haustoria via α-linked mannoside residues on the haustorial surface which are common in glycoproteins and wall polysaccharides of rust fungi (Kim et al., 1982, Hahn and Mendgen, 1992). Infected leaves 9 dai (one day prior to sporulation) were collected and blended using a Waring blender, filtered sequentially through nylon meshes of pore size 100 µm and 20 µm, and loaded onto a column containing ConA-Sepharose beads. Haustoria retained on the column were released by mechanical agitation and collected immediately by retaining the flowthrough once the beads had settled. Light microscopy showed that this protocol mainly yielded a mixture of single-cell haustoria and contaminating intact and broken chloroplasts (Figure 2.3)

Figure 2.3. Stripe rust haustoria isolated by affinity chromatography. Haustoria isolated by ConA affinity chromatography on homogenized extracts that were prepared with a 20 µm mesh to remove plant cell debris. The black arrows show haustoria and the white arrows indicate contaminating chloroplasts. Bright field image collected on a Leica DMR epifluorescence microscope.
Most isolated haustoria had retained their neck, which is thought to sealing the extrahaustorial matrix from the plant apoplast, both facilitating nutrient uptake and connecting the haustorium with the rest of the fungal body (Voegle and Mendgen, 2003) (Figure 2.4). Many haustoria showed elongated irregularly lobed shapes. ConA conjugated to the fluorophore Alexa-fluor 594 (excitation/emission peaks at 590/617 nm), or the lectin wheat germ agglutinin (WGA) conjugated to the fluorophore fluorescein isothiocyanate (FITC, excitation/emission peaks at 495/517 nm), were used to assess the purity and integrity of isolated haustoria (Figure 2.4). WGA is a lectin that binds specifically to trimers of N-acetyl-D-glucosamine which is the structural unit of chitin (Goldstein and Poretz, 1986), the major cell wall component of many fungal species. The fluorescent images suggest that haustoria were isolated as intact single-cell structures, while bright field images showed that most of the haustoria retained their cytoplasmic contents, making them suitable for experimentation. Although both lectins specifically bound haustoria and not contaminating chloroplasts, WGA did not bind haustoria as consistently as ConA, as not all haustoria in a single field were stained with the same intensity (Figure 2.4).

Figure 2.4. Staining of isolated *Pst* haustoria with ConA and WGA. Bright field and fluorescent images showing isolated haustoria after 30 min incubation with WGA-FITC or ConA-Alexa 594. Black arrows indicate the presence of a visible neck and white arrows show young haustoria. All images were collected on a Leica DMR epifluorescence microscope, 40x magnification.

### 2.3.3. RNA isolation from haustoria and spores and purification of mRNA
Extraction of total RNA from isolated haustoria and purification of mRNA were the next steps in obtaining material for transcriptome sequencing. A minimum of 200 ng of high quality mRNA was necessary for the preparation of a cDNA library for 454 RNA-seq. In eukaryotic cells, mRNA comprises approximately 1–5% of total cellular RNA, although the actual amount depends on the type of cell and its physiological state (Bryant and Manning, 2000, Warner, 1999). Thus, a minimum of 16 µg of total RNA from haustoria samples was calculated to be necessary for mRNA purification. On average, 2 µg of total RNA was obtained from 15 g of Pst-infected wheat tissue, so a total of ~120 g of infected tissue was processed in 8 procedures.

Messenger RNA was purified from total RNA based on capture of the poly A tail by oligo dT covalently coupled to magnetic beads. The sizing, quantification and quality control of the isolated mRNA was assessed on the Agilent 2100 Bioanalyzer, using the RNA 6000 pico kit (Kuschel, 2000). Figure 2.5 shows the electronic gel-like image and the electropherogram obtained from the extracted mRNA from haustoria tissue.

Figure 2.5. mRNA analysis with the Agilent 2100 bioanalyzer. A. Electronic gel-like image generated for the RNA 6000 ladder (M) and haustorial mRNA (H). B. "Live" plot of the migration time against fluorescence units of nucleic acids contained in haustorial mRNA preparation. The asterisk represents the peak typical of ribosomal contamination.

The electropherogram indicated that the pattern obtained was typical of high quality mRNA (Kuschel, 2000), containing 4% ribosomal contamination. This amount is considered low. The total amount of mRNA was ~210 ng, sufficient for preparation of the cDNA library, and contamination with genomic DNA was not detected as slow migrating peaks were absent.
Total RNA and mRNA were also extracted from spores following the same methods as for haustoria samples, although RT-PCR tests were not performed as this tissue by its nature is free of contaminating plant material. The Bioanalyzer was used to verify the concentration and quality of the mRNA. mRNA samples were submitted to the Biomolecular Resource Facility at ANU who performed cDNA library preparation, sequencing and data quality control. Newbler, the dedicated software package for de novo DNA sequence assembly for data generated on the 454 platform, was used only to remove poor quality reads and the adapter sequences (short pieces of DNA which were ligated to the end of cDNA fragments to prime the sequencing primers).

2.3.4. Assembly of contigs from haustoria and spore RNA-seq data, and prediction of effector genes

RNA-seq, also called whole transcriptome shotgun sequencing, is the use of high-throughput cDNA sequencing to derive information about the original RNA sample (Wang et al., 2009c). Two of the most common RNA-seq technology platforms are the Illumina Genome Analyzer (Bentley et al., 2008) and 454 (Margulies et al., 2005). There are a number of differences between these platforms in terms of technology which influence run time, sample preparation, amount of data generated, and total cost of the procedure. Illumina reaches high coverage per run at lower costs (Zagordi et al., 2012), whereas 454 produces much longer reads than Illumina which facilitates the assembly process. As there was no Pst genomic or transcriptomic data available at the time this project started, I reasoned that the use of 454 would be a better option to produce transcriptomic data as the long reads would be more easily assembled de novo. Later in this project (Chapter 3), Illumina RNA-seq data was also produced to complement the 454 analysis.

Messenger RNA samples from enriched stripe rust haustoria and germinated spores were sequenced by single-read pyrosequencing on a Roche 454 GS-FLX titanium instrument. A total of 301 Mb and 191 Mb of sequence data were generated in the form of 729,036 and 457,071 reads averaging 413 bp and 420 bp in length for haustoria and spores, respectively (raw data deposited in Sequence Read Archive under accession numbers SRR579539 and SRR579540). Approximately 1.18% and less than 1% of the reads were discarded after quality control filters were applied to the sequencing data from haustoria and spores respectively. To assemble the spore reads de novo, CLC genomics (CLC Bio 3.9) was used with optimized parameters: similarity 0.97 and overlap 0.5, producing 7,886 contigs. The number of reads per contig ranged between two and 27,778, the level of coverage ranged from one to 23,719, and the size ranged between 200 bp and 4765 bp with an average length of 794 bp.
(Figures 2.6-2.8). Haustoria isolated by affinity chromatography contained a considerable proportion of contaminating chloroplasts, and although RT-PCR tests indicated no significant contamination with plant mRNA, it was expected that a small percentage of the raw reads from the haustoria sample would be of plant origin. To remove contaminating wheat sequences from the haustoria data, the draft genome of the American stripe rust isolate Pst-130 (Cantu et al., 2011), a draft genome of Pst 104 E137 A- (Jackson et al., unpublished data; Chapter 6) and the de novo assembled transcriptome from spores (Transcriptome Shotgun Assembly project deposited at DDBJ/EMBL/GenBank under the accession GAIIY00000000) were used as references to extract fungal sequences which accounted for approximately 62% of the original reads. Since these references do not represent the complete Pst genome, the remaining unmapped reads were assembled de novo and the resulting contigs screened by BLAST search against the NCBI nucleotide and protein databases. The BLAST result was curated manually and contigs showing hits to plant genes were removed, while the remaining contigs were retained as novel transcripts not included in the draft genome assemblies. The filtered haustorial reads were then assembled de novo using CLC genomics (CLC Bio v3.9) resulting in 12,846 contigs representing the haustorial transcriptome (Transcriptome Shotgun Assembly deposited at the DDBJ/EMBL/GenBank under the accession GAIS00000000). Assembled contigs ranged between 200 and 6,854 bp, with an average length of 704 bp (Figures 2.6 and 2.7). Most of the contigs in both datasets (69.14% and 74.33% in haustoria and spores respectively) are bigger than 500 bp (Figure 2.7), showing that 454 platform is useful for the generation of larger contigs, increasing the probability of predicting full-length open reading frames (ORFs).

Figure 2.6. Number of 454 reads per assembled contig for spores and haustorial samples.
To identify candidate effector genes, haustorial contigs were analysed using CLC genomics to identify all possible ORFs, which were further analyzed with SignalP v3.0 and v4.0 (Bendtsen et al., 2004) to predict secretion signal peptides (SP). A total of 1,299 SP-encoding genes were found and filtered to include only those unique to a contig, or alternatively the largest SP-ORF within a contig which is most probably the reading frame that will be translated into the protein (Kumar et al., 2007). Proteins containing transmembrane domains (predicted using TMHMM 2.0 (Krogh et al., 2001)) or mitochondrial targeting signals (predicted using TargetP (Emanuelsson et al., 2007)) were excluded.
This left 437 candidate effector genes, from now onward called predicted haustorial secreted proteins (HSP) (Table S2 and Data S3 in Garnica et al., 2013).

2.3.5. Validation of expression of HSP genes

To validate the in silico gene predictions and confirm the fungal origin of the effector candidates, 94 out of the 437 HSP genes were chosen randomly and tested for expression in vivo (Figure 2.9). Primers were designed to amplify the full ORF excluding the SP-encoding sequence. To test expression, cDNA samples were prepared from spores, isolated haustoria, wheat tissue 9 dai, and uninfected wheat tissue, and subjected to RT-PCR using HSP-specific primers (Appendix 1). Pst genomic DNA was also included as a secondary control to confirm the fungal origin of the tested genes. Primer pairs for Pst β-tubulin or wheat specific genes as glyceraldehyde 3-phosphate dehydrogenase, PR2 (accession Y18212) and PR5 (accession AF442967) served as positive controls (Appendix 1).

Figure 2.9. HSPs expression patterns studied by RT-PCR. Ninety four HSPs were randomly selected and their expression patterns were analyzed by RT-PCR. Primers were designed to amplify the full length gene sequence minus the signal peptide-encoding region. Pst 104 E137 A- genomic DNA, and cDNAs from spores, isolated haustoria, infected wheat leaves, and uninfected wheat tissue were used as templates. Two expression patterns were obtained: 1. Expression only during the biotrophic phase (Pstv_3161-1), and 2. Expression during early and late stages of development (Pstv_7145-1). None of the tested effector candidates were amplified from uninfected wheat leaves.
Seventy one genes showed expression in isolated haustoria and infected wheat tissue, but no detectable expression in spores; 22 showed expression in all tissues where the fungus was present; and one failed to amplify. Ninety one of these genes were amplified from genomic DNA and/or cDNA from spores, confirming their fungal origin. The remaining three genes were also confirmed to be of fungal origin as they were identified in our draft genome of this isolate (Jackson et al., unpublished data) but were not amplified from genomic DNA because one or both primers span intronic regions. Cloning and Sanger sequencing of 25 candidate effectors confirmed that they corresponded to the sequences predicted from my 454 assemblies.

### 2.3.6. Expression profiles of HSP genes

To better estimate the levels of expression of the 437 previously predicted HSPs in each fungal tissue, I used digital gene expression analysis based on data generated on the Illumina RNA-seq platform (Illumina sequencing data are broadly described and analyzed in Chapter 3) (Table S2 in Garnica et al., 2013). Illumina sequencing provided considerably more reads and much deeper coverage, allowing me to better estimate the level of expression of the 437 HSPs, and better statistical comparisons between pathogenic stages. Three biological replicates per pathogenic stage (haustoria and germinated spores) were sequenced and the reads obtained for each sample were mapped against the HSP nucleotide sequences. A total of 21.5 million reads were mapped against the HSP genes. The digital expression analysis revealed the average level of expression of each predicted HSP in haustoria and spores. All tested genes in the previous section showed a very close correlation between the expression pattern detected with RT-PCR, and that predicted in silico (Table S2 in Garnica et al., 2013). Remarkably 371 (85%) of the HSP genes were expressed differentially; 295 were overexpressed in haustoria, and 76 were overexpressed in spores. Strikingly, 140 (40%) of the HSPs overexpressed in haustoria showed no expression in spores and 156 (50%) were ten times or more expressed in haustoria than in spores (Figure 2.10).
2.3.7. Annotation of HSP genes

2.3.7.1. Length and cysteine content

Extracellular proteins frequently contain elevated numbers of cysteine residues, which can participate in di-sulfide bonding necessary for their folding and stability in the protease-rich apoplastic space. Many known and predicted effector proteins from filamentous pathogens are small cysteine-rich proteins (Stergiopoulos and de Wit, 2009, Saunders et al., 2012, Hacquard et al., 2012). Although the presence of a high number of cysteine residues is not a criterion to predict and/or select effector candidates, many studies have shown that the cysteine residues present in a number of confirmed effector proteins plays key roles in structure and activity (Kooman-Gersmann et al., 1998, Stergiopoulos et al., 2007, Catanzariti et al., 2006). Analysis of the cysteine content of the predicted 437 HSP proteins revealed that 246 contained fewer than four, and 191 contained from 4-28 cysteine residues in the mature protein (hence designated Cys-rich). Most of these (~70%) were expressed preferentially in haustoria (Table S2 in Garnica et al., 2013).

Analysis of conserved cysteine spacing of the Cys-rich predicted proteins was done in collaboration with Dr Brendon Conlan (Rathjen Laboratory, RSB, ANU). The predicted amino acid (aa) sequences of the Cys-rich HSPs were first clustered according to the total number of cysteines. Manual analyses

Figure 2.10. HSPs expression patterns. Number of HSPs showing digital expression patterns as shown. Overexpression was evaluated using Baggerley’s test (Baggerly et al., 2003), genes with a false discovery rate corrected p-value less than 0.05, fold change >2 and a difference of at least 20 were considered to be significant.
were done in Microsoft Excel to detect cysteine residues located in the same position within different HSP aa sequences, and then pairs of cysteine residues with identical or very similar spacing (the number of aa between cysteine residues). A total of 88 genes could be grouped into 21 classes in relation to the presence of a defined number of cysteine residues with particular spacing arrangements between them. Table 2.1 shows the cysteine pattern corresponding to each class, and Figure 2.11 represents the location of each motif within the predicted mature HSP proteins.

<table>
<thead>
<tr>
<th>Class</th>
<th>Number of genes</th>
<th>Cys Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>37</td>
<td>C(6-15)-C(9-21)-C(6-22)-C-C</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>C(16-11)-C(4)-C(5-7)-C(2)-C(13-25)-C(12-17)-C(8-9)-C(3)-C(5)-C(2)-C-C</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>C(2-4)-C(12-14)-C(4-5)-C(2-4)-C(10-11)-C(2-3)-C</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>C(11-12)-C(7-9)-C(3)-C(8-9)-C(14-16)-C(13-15)-C(8-9)-C(9-11)-C(2)-C</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>C(4)-C(5)-C(26)-28-29-C</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>C(13)-C(7-8)-C(13)-38-C</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>C(14)-C(12)-C(19)-C(11)-C</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>C(1)-C(9-11)-C(18)-C(5-13)-C(12-29)-C(32)-C(1)-C(9-11)-C(18-19)-C(5)-C</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>C(2)-C(4)-C(11-12)-C(2)-C(7)-C(10-12)-C(2)-C-C</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>C(18)-C(3-4)-C(3-4)-C</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>C(21-22)-C(8-93)-C(3)-C(8-10)-C(8)-C(15-17)-C(37-43)-C(71)-C</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>C(14)-C(27)-C(4)-C(15)-C(28)-C(9)-C(44)-C(2)-C</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>C(10)-C(28)-C(27)-C(40)-C(13)-C</td>
</tr>
<tr>
<td>14</td>
<td>2</td>
<td>C(11)-C(8-11)-C(4)-C(14-25)-C(33-43)-C(13-14)-C(2)-C(1)-C</td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>C(4)-C(9)-C(4)-C(3)-C(5)-C</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>C(48-50)-C(11)-C(2)-28-C(20)-C(13-14)-C</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>C(12)-C(24)-C(2)-C(32)-C(28-27)-C</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>C(13)-C(4)-C(12)-C(2)-C(12)-C</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>C(42-45)-C(14)-C(19)-C(31)-C</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>C(4)-2-C(3)-C(2)-C</td>
</tr>
<tr>
<td>21</td>
<td>2</td>
<td>C(54)-C(15)-C(9)-C</td>
</tr>
<tr>
<td>Total</td>
<td>88</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1. Pst HSP classes with characteristic Cys patterns. Capital “C” represents a cysteine residue, lower case “x” represents any aa and the numbers in brackets represent the number of possible residues between two cysteine residues. The colour key for each class numbers corresponds with the protein structure diagram of Figure 2.11.
Figure 2.11. Effector candidates grouped according to distribution of cysteines in the predicted protein sequence. The horizontal white bars represent the total lengths of the predicted HSP aa sequences (without the signal peptide) and the superimposed coloured bars represent the length and location of the Cys patterns described in Table 2.1. The colour key of the bars corresponds with the colour key in Table 2.1.
Although none of the cysteine patterns found in Pst HSPs matched cysteine spacing motifs found in candidate effectors from other rust fungi (Hacquard et al., 2012), further analyses are necessary to determine if some of the Pst Cys spacing motifs fit those characteristic of known enzyme families (Cheek et al., 2006).

2.3.7.2. Functional classification

In an attempt to predict functions for the HSP genes, I sought sequence similarities to proteins of documented function, and presence of conserved protein motifs, using BLAST and BLAST2GO. The predicted mature peptide sequences of the Pst HSPs (minus the SP) were searched against the NCBI non-redundant protein database using BLASTx. They were additionally analyzed with BLAST2GO (B2G) (Conesa et al., 2005) to find gene ontology terms associated with the HSPs, and the presence
of conserved protein signatures through the tool InterProScan. Approximately 263 of these proteins had no similarity to known proteins at the level of e-val >10^{-25}. One hundred and five proteins were similar to hypothetical proteins from Pgt, and four were similar to previously identified Pst secreted proteins (Dong et al., 2011). Forty HSPs could be partially annotated with B2G, fifteen of which were classified as glycoside hydrolases from different families (Cantarel et al., 2009), especially family 18 (chitinase activity) and families related to plant cell wall degradation. Four HSPs were annotated as putative polysaccharide deacetylases (chitin deacetylases), and the rest were annotated as putative proteins participating in diverse cellular processes such as protein folding, proteolysis, oxidation-reduction, and regulation of transcription (Table S2 in Garnica et al., 2013). A further 35 HSPs contained conserved protein domains such as zinc finger domain, copper/zinc binding domain, FAD-binding domain, cupredoxin domain, barwin-like endoglucanase domain, NUDIX hydrolase domain, thaumatin, phosphoryltryphanolamine-binding domain, CFEM domain and MD-2-related lipid-recognition domain. Most of these roles or domains have been identified in other rust predicted secretomes (Duplessis et al., 2011, Saunders et al., 2012), suggesting conserved roles in Pucciniales. Interestingly, a novel effector gene recently reported in Colletotrichum truncatum (CtNUDIX) is expressed exclusively during the late biotrophic phase and elicits the hypersensitive response in tobacco leaves, suggesting that this effector could be important for the transition from biotrophy to necrotrophy, or alternatively is recognized by the tobacco immune system (Bhadauria et al., 2012). A secreted protein from Pst containing a NUDIX domain could potentially regulate redox homeostasis. Nevertheless, as observed previously for rust fungi, the absence of recognisable protein domains in candidate effector proteins is a common finding (Saunders et al., 2012).

Genome and transcriptome sequencing of Blumeria graminis revealed that the most highly expressed candidate effector proteins contained a Y/F/WxC motif (tyrosine /phenylalanine /tryptophan /any aa/ cysteine) 1-30 aa after the predicted SP (Spanu et al., 2010). I identified 1 to 4 copies of this sequence motif in 124 of the 437 Pst HSPs, but the motif occurred within the first 30 aa for only 43 of these, which is similar to the number expected by chance (~32). Thus similar to Pgt and Mlp (Duplessis et al., 2011), this motif does not seem to define a major class of effectors in Pst. Further attempts to detect novel motifs using the MEME analysis module within CLC failed to identify other conserved sequences.

2.3.8. Potential recognition of Pst HSPs by wheat tester lines
One informative approach to validate the effector nature for the predicted \textit{Pst} \textit{HSP} genes is to investigate if they can be recognized by stripe rust resistance genes (\textit{R} genes, also called \textit{Yr} genes when they are specific for stripe rust (McIntosh et al., 1995)) present in wheat. This is based on the assumption that effectors of rust fungi are translocated into the host cytoplasm (Rafiqi et al., 2010, Kemen et al., 2005) where they modify or perturb host cellular processes (Schornack et al., 2010, Kamoun, 2007, Hogenhout et al., 2009). This possibility was tested for 23 \textit{Pst} \textit{HSP} genes, using a bacterial delivery system on wheat cultivars carrying different \textit{R} genes known to recognize \textit{Pst} 104 E137 A- (Wellings, 2007). The 23 \textit{HSP} genes were chosen on the basis of their high relative expression (from 454 RNA-seq reads, data not shown), small size (encoding proteins <300 aa), pattern of gene expression (expressed only in haustoria and infected tissue), and the absence of homology to known genes (Table S2 in Garnica et al., 2013).

The TTSS is essential for many gram-negative bacteria to infect their animal or plant hosts, and serves to translocate bacterial proteins into eukaryotic host cells to manipulate them during infection (Blocker et al., 2003). The bacterial system used for this study was the engineered \textit{P. fluorescens} Pf0-1 EthAn (Effector-to-Host Analyzer), which is effective for delivering type-III effector proteins into host cells (Thomas et al., 2009). Recently this system has been used successfully to deliver rust effector proteins into wheat cells (Upadhyaya et al., 2014, Yin and Hulbert, 2010). Here, I amplified \textit{HSP} ORFs lacking the SP-encoding sequence (no-SP ORFs) of selected \textit{HSP}s genes from haustorial cDNA, and cloned them into the bacterial expression vector pEDV-AvrRPM1-G2AC3A-Des1 (Figure 2.12) (Upadhyaya et al., 2014) using the Gateway recombination system. The cloning site of this vector allows fusion of the gene of interest with a sequence encoding the TTSS secretion signal (SS) of the bacterial avirulence gene \textit{AvrRPM1}, which allows translocation of proteins into the host cell through the TTSS (Grant et al., 2006). The myristoylation (Gly-2) and palmitoylation (Cys-3) residues naturally encoded in the secretion-translocation signal of \textit{avrRPM1} were mutated to alanine residues (Upadhyaya et al., 2014), to avoid localization of the recombinant proteins to the host plasma membrane via post-translational modification with the fatty acids myristate and palmitate, respectively (Nimchuk et al., 2000). In this system, the \textit{AvrRPM1(SS)-HSP} fusions are controlled by the \textit{AvrRPM1} promoter which is itself induced by the pETHAN type-III system \textit{in planta} or in minimal media at low pH (Ritter and Dangl, 1995)(Figure 2.12).

The effector candidates cloned into the expression vector were sequenced using the Sanger method (Sanger et al., 1977) and transformed into \textit{P. fluorescens} Pf0-1 (EthAn) by tripartite mating. Positive colonies were passaged through single colonies twice, and were confirmed by PCR for each \textit{HSP} candidate as well as for the \textit{P. fluorescens} 16S ribosomal subunit.
Figure 2.12. Schematic representation for delivery of Pst HSPs via the TTSS of P. fluorescens Pf0-1 (EtHAn). The Gateway compatible vector pEDV-AvrRPM1-G2AC3A-Des1 enables cloning of the HSP genes downstream of the AvrRPM1 promoter and AvrRPM1(SS). ATG and STOP indicates the initiation and stop codons for the fusion genes. The arrows indicate the mutation of the Gly-2 and Cys-5 codons on the AvrRPM1(SS). The asterisk represents a plant immune receptor. ER: Endoplasmic reticulum of the plant cell. The AvrRPM1 promoter induces the expression of the AvrRMP1(SS)-HSP fusions in response to the environmental conditions of the plant apoplast. AvrRPM1(SS)-HSP fusions are then synthesized in P. fluorescens and translocated into the plant cell cytoplasm where they can hypothetically travel to different subcellular locations for different functions.

To select the wheat cultivars to screen effector recognition, the Pst strain 104 E137 A- was tested for pathogenicity on the original wheat differential set used by Prof. Robert McIntosh (Plant Breeding Institute, Cobbity NSW) for its designation. This set contains 18 cultivars, 15 from the differential set proposed by (Johnson et al., 1972) and three added later for Australian studies of Pst (Wellings and McIntosh, 1990) (Table 2.2). Infections on this differential set confirmed the identity of the Pst strain, and a total of 10 wheat cultivars which were resistant to infection were selected for the ability to recognize Pst HSPs delivered by bacterial TTSS.
Table 2.2. Pathotyping Pst strain 104 E137 A- to select resistant wheat cultivars. This table shows the eighteen wheat cultivars routinely used to designate Australian Pst strains. The expected and obtained resistance (R) or susceptibility (S) reactions when these cultivars are infected with Pst 104 E137 A- are shown. The cultivars in yellow were tested for their ability to recognize Pst HSPs delivered by the bacterial TTSS.

Seeds from the resistant cultivars were vernalized and germinated in petri dishes to coordinate their germination. Germinated seeds were transferred to soil and grown for 10 days or until the second leaf was adequately expanded for infiltration. Expression of TTSS genes in P. fluorescens Pf0-1 (EtHAn) carrying each Pst HSP was activated by maintaining the bacteria in minimal media (pH 5.7) overnight at 20°C, and then bacteria were pelleted, resuspended in infiltration buffer (10 mM MgCl₂) to an OD₆₀₀ of 1.0 and infiltrated immediately into wheat leaves. Three plants per wheat genotype were infiltrated for each HSP gene, and the presence or absence of HR was scored 5 and 8 days after infiltration. Every effector was tested against every wheat cultivar three times. At the time that the experiments were done, no Pst or Pgt Avr genes that could be used as positive controls had been cloned, so the genes AvrM and M from the flax rust-flax pathosystem (Catanzariti et al., 2006) were used as a positive control for activation and function of the bacterial TTSS. P. fluorescens carrying AvrM cloned in pEDV-AvrRPM1-G2AC3A-Des1 was grown and activated together with the Pst HSP, and was infiltrated in transgenic tobacco plants expressing M (Catanzariti et al., 2006) (Figure 2.13).
Figure 2.13. *AvrM-M* recognition in tobacco. Necrosis produced in transgenic *M* tobacco plants two days after infiltration with *P. fluorescens* carrying *AvrM* (1) or the inactive form *avrM* (2) fused to the *AvrRPM1*(SS). Transient genetic transformation of tobacco leaves was performed by infiltrating *Agrobacterium tumefaciens* carrying carrying 35S:*AvrM* (3) or 35S:*avrM* (4) to compare the effectiveness of TTSS delivery and transient expression. Infiltration buffer was used as an infiltration control (5).

None of the 23 tested *Pst* HSPs caused necrosis on any of the wheat cultivars used for this assay. Although the *P. fluorescens* method is a convenient way to introduce the HSP proteins into the host cytoplasm, the lack of time allowed me to test only a small number of HSPs. The long process of cloning HSPs into the pEDV-AvrRPM1-G2AC3A-Des1 vector and its transformation into *P. fluorescens* was a limiting step that prevented me testing more candidates. However, this approach is currently being extended and improved in this laboratory by (i) cloning a larger number of candidate HSP genes, (ii) including non-host grass species in the set of hosts and (iii) using a modified version of the pEDV-AvrRPM1-G2AC3A-Des1 vector which includes epitope tags downstream of the HSPs to track the fusion proteins by western blot.

### 2.3.9. Assays for inhibition of plant cell death (PCD) by *Pst* HSPs.

The second approach to characterize *Pst* HSPs was to test their capacity to inhibit plant cell death. Biotrophic fungal pathogens colonize living plant tissue and obtain nutrients only from living host cells, therefore it is important for them to suppress the hypersensitive response. Thus, it has been hypothesized that one of the roles of effector proteins is to act as suppressors of cell death (Panstruga, 2003). To investigate this hypothesis with *Pst*, I developed a new assay based on specific elicitation of cell death by the fungal ToxA protein (Friesen et al., 2006). ToxA is a toxin produced by
the wheat necrotrophic pathogens Stagonospora nodorum and Pyrenophora tritici-repentis (Friesen et al., 2006), and plays a major role in production of necrotic lesions (dead cells) surrounding the infection sites.

I hypothesized that if Pst effectors are capable of suppressing cell death, plant tissue colonized by Pst would be attenuated in the necrotic response to ToxA. To test this, the wheat cultivars B26 and Grandin, which contain the susceptibility gene Tsn1 necessary for ToxA activity, were infected with Pst 104 E137 A- to determine if they were susceptible to this strain. Cultivar Grandin showed the greatest susceptibility to Pst 104 E137 A- (data not shown), so was used for the plant cell death inhibition assay. I first purified the toxin in collaboration with Dr. Brendon Conlan (Rathjen Laboratory, RSB, ANU). To test the ToxA activity of the purified toxin and determine the optimal concentration for the plant cell death inhibition assays, serial dilutions of the concentrated protein were syringe infiltrated into leaves of 7 days old seedlings of wheat cv. Grandin, and chlorosis and necrosis symptoms were scored at 2, 3 and 5 days after infiltration, Figure 2.14.

I selected the ToxA dilution of 1:30 which gave severe chlorosis two days after infiltration, and necrosis three days after infiltration, as a convenient sequence of events for development of the plant cell death inhibition assay.
In the first assay to test suppression of ToxA cell death, I infected seven day old cv. Grandin seedlings with *Pst* and treated the infected leaves with diluted ToxA (1:30) 7 dai. The appearance of characteristic ToxA symptoms with respect to uninfected controls was recorded daily, and pictures were taken four days after infiltration (the first day after *Pst* sporulation) as shown in Figure 2.15.

![Figure 2.15](image)

**Figure 2.15. Inhibition of ToxA-induced plant cell death by *Pst* infection.** Infected and uninfected cv. Grandin plants were infiltrated with purified ToxA at a dilution of 1:30, and negative controls were infiltrated with water. All plants were maintained under the same growing and infection temperature throughout the experiment, and infiltrations were done early in the morning as a minimum 9 h of light is required for ToxA activity (Manning and Ciuffetti, 2005). This experiment was done three times with similar results.

Although infected and uninfected plants were maintained at 17°C for the whole experiment (the optimal temperature for *Pst* growth), strong chlorotic and necrotic symptoms induced by ToxA on the uninfected control appeared at 2 and 3 days after infiltration respectively, suggesting that the timing of ToxA-induced effects is the same at lower temperatures. In tissue infected with *Pst* 104 E137 A-, necrosis was notably reduced in areas where *Pst* pustules were visible. Importantly, the fungus was able to produce spores 9 dai, the same time point as for the controls infiltrated with water, in those places where the toxin was infiltrated. The data suggest that the fungus could continue its normal growth process despite the presence of ToxA, possibly by suppressing the action of ToxA in some way.

Thus, since the presence of *Pst* apparently suppresses the ToxA effect to some extent, one possibility could be that individual effectors translocated to the host cytoplasm inhibit the ToxA-induced cell death. To investigate this hypothesis, *P. fluorescens* was used to deliver effector candidates to the
host cytoplasm, and 16 h later the same leaf areas were infiltrated with ToxA. Chlorosis and necrosis were scored in relation to the controls at 24, 48 and 72 h after infiltration on an arbitrary scale of 1 to 5, where 1 is completely healthy tissue, 2 is chlorotic tissue, 3 is slightly necrosed chlorotic tissue, 4 is notably necrosed with some remaining chlorosis, and 5 is completely necrosed tissue (Figure 2.16A). Three days after infiltration leaves were detached and pictures were taken (Figure 2.16B). As a negative control for cell death inhibition, *P. fluorescens* carrying an empty vector construct (EV) was infiltrated into leaf areas previously infiltrated with ToxA. As a further negative control, *P. fluorescens* carrying each *HSP* gene was infiltrated with water, to account for intrinsic necrosis caused by any of the *HSPs*. The positive cell death control was ToxA plus infiltration buffer. It was not possible to include a positive control of the inhibition of plant cell death as to date no such inhibitor is known.

A

![Graph A](image)

B

![Graph B](image)
Chapter 2: Prediction and characterization of Pst effectors

<table>
<thead>
<tr>
<th>P. fluorescens EV</th>
<th>Infiltration buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>water</td>
<td>ToxA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HSP 6</th>
<th>HSP 7</th>
<th>HSP 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>water</td>
<td>ToxA</td>
<td>water</td>
</tr>
</tbody>
</table>

[Image of leaf samples showing different treatments with P. fluorescens EV and infiltration buffer, alongside HSP 6, HSP 7, and HSP 12 treatments.]
Figure 2.16. Inhibition of ToxA-induced cell death by bacterial delivery of Pst HSPs. Wheat cv. Grandin seedlings were infiltrated with *P. fluorescens* carrying each HSP gene, and subsequently with ToxA. A. The average score (three experiments) of symptoms developed 3 days after ToxA infiltration were recorded for each HSP and controls, and results were plotted. Standard deviation bars are indicated. B. Examples of results obtained with some HSPs are shown. Pictures were taken 72 h after ToxA infiltration. Negative controls comprise infiltration of water instead of ToxA. The numbers on the pictures relate to their original IDs in Garnica et al., 2013 as follow: HSP 1 = Pstv_3046, HSP 2 = Pstv_3060, HSP 3 = Pstv_3161-1, HSP 4 = Pstv_3271, HSP 5 = Pstv_4624-1, HSP 6 = Pstv_4790, HSP 7 = Pstv_4823, HSP 8 = Pstv_4852, HSP 9 = Pstv_4853, HSP 10 = Pstv_4902-1, HSP 11 = Pstv_5075, HSP 12 = Pstv_5088-1, HSP 13 = Pstv_5365, HSP 14 = Pstv_9199, HSP 15 = Pstv_12358, HSP 16 = Pstv_12461, HSP 17 = Pstv_12696, HSP 18 = Pstv_12697, HSP 19 = Pstv_13351, HSP 20 = Pstv_13398, HSP 21 = Pstv_13424, HSP 22 = Pstv_15884, HSP 23 = Pstv_16344.

Although different periods between the times of bacterial and ToxA infiltration were tested, 16 h appeared to give the clearest results (data not shown). Possibly, this gives sufficient time for the bacteria to synthesize and translocate the HSP proteins into the host cytoplasm, so that they can exert their function prior to the induction of necrosis by ToxA.

Leaves infiltrated with ToxA only were severely chlorotic after 24 h, and completely necrosed three days after infiltration. Plants infiltrated with bacteria containing EV and ToxA showed necrosis three days after infection, but not to the same extent as when ToxA was infiltrated by itself or with infiltration buffer (Figure 2.16). Similarly, none of the leaves co-infiltrated with HSPs and ToxA died at the same speed as when ToxA was infiltrated alone. This suggests that the presence of bacteria delays necrosis. Although this background suppression made scoring more difficult, HSP proteins #6,
#7, and #12 clearly showed less necrosis and low values on the scoring scale. These three HSPs consistently showed a pronounced delay in development of chlorosis and necrosis, in all repeats of the experiment (Figure 2.16). This suggests that these three effector candidates might inhibit or impair the responses induced by ToxA.

2.4. DISCUSSION

Affinity chromatography of isolated Pst haustoria from infected wheat tissue combined with 454-pyrosequencing allowed me to produce the first next-generation Pst haustoria transcriptome. This transcriptome was the basis for prediction of putative Pst effector genes which are expressed in haustoria, the most important pathogen-host interface during infection. Additional tools such as RT-PCR and a newly developed technique to isolate haustoria (discussed further in Chapter 4), combined with illumina sequencing of haustoria and germinated spores helped to confirm the expression profile of the whole set of predicted effector genes. This suggested that the expression of effector genes in the haustorium is very active as has been proposed previously (Catanzariti et al., 2006). However, contrary to what was expected, 17% of these genes were expressed at high levels in germinating spores, indicating that this pathogenic stage could also play an important role in effector synthesis and delivery. Accordingly, the expression and deployment of effectors from other pathogenic stages such as infectious hyphae, other spore types, or other pathogenesis stages (i.e. infections on a secondary host) is also expected and could differ from the current observations. Among the 437 effector candidate genes, 368 had no significant similarity to known genes or were similar to putative proteins of unknown function, revealing that the vast majority of them lacked traits that could indicate their roles during pathogenesis.

2.4.1. Purification of Pst haustoria from infected leaf tissue

Using the ConA affinity chromatography technique (Hahn and Mendgen, 1992), isolation of Pst haustoria was possible after optimising the method for this species. One of the critical steps was the filtration of tissue extracts through nylon meshes to remove cell debris. Preliminary tests with nylon meshes of different pore sizes (11 µm, 15 µm, 20 µm and 40 µm) for the second filtration step suggested that 20 µm was the best size to avoid contamination with large fragments of plant cell debris. Also, the use of this mesh size did not appear to generate a bias for small immature haustoria as the isolated haustoria ranged in size and shape, including the small spherical immature haustoria, and the apically branched ones characteristic of mature haustoria (Sorensen et al., 2012). Repeated
isolations using a nylon mesh of 20 µm consistently showed a high number of haustoria compared to very low amounts of other contaminants apart from chloroplasts, as for example fungal hyphae, intact plant cells, plant cell debris, or fungal spores. Due to the proposed use of this material in the preparation of cDNA libraries for 454 –FLX pyrosequencing, the removal of contaminating chloroplasts was not considered essential as previous studies have shown that the total RNA from these organelles is insignificant (Hahn and Mendgen, 1997, Catanzariti et al., 2006). Additionally, the absence of polyadenylation in RNA from chloroplasts (Mayfield et al., 1995) means that it will not be transcribed to cDNA by the Oligo dT method used in this study.

Haemocytometer counts were often performed to determine the efficiency of the isolation technique, and on average, $6.0 \times 10^5$ haustoria were isolated per 15 g of heavily infected tissue with a haustoria:chloroplast ratio of approximately 1:15. Similar numbers have been reported when this method has been applied to the rust fungus *Melampsora lini* (Catanzariti, 2006), although about 60-fold more haustoria were obtained from *Uromyces vicia-fabae* (Hahn and Mendgen, 1992). These differences in yield could be caused in part by structural differences between pathosystems (differences in the content or exposure of the specific sugar residues to which ConA binds), as well as by the yield sacrifices made in order to optimize the balance between haustoria recovery and reduction of contamination. Due to the very low efficiency of this method when applied to *Pst*, several rounds of haustoria isolation were necessary to obtain sufficient mRNA for 454 sequencing. Thus, alternative ways to enrich haustoria from infected tissue were explored, and a newly developed technique is discussed in Chapter 4.

Bright field and fluorescent microscope images suggested that the ConA-enriched haustoria retained their structural integrity. Interestingly, although both lectins specifically bound haustoria and not contaminating chloroplasts, they appeared to have different affinities. While ConA stained all haustoria evenly, WGA did not show consistent staining. Previous studies on the wall composition of wheat rusts using different lectins have revealed contradictory results. The model described for *Pgt* by Chong et al. (1985, 1986) predicts that WGA binding sites on isolated haustoria would only be detected if the extrahaustorial membrane and matrix were lost during purification, and that in any case, no labeling of the neck area is ever observed (Chong et al., 1986, Chong et al., 1985). Alternatively, ConA binds the haustorial neck, the haustorial wall, and extrahaustorial matrix in young haustoria, but only the extrahaustorial matrix in older haustoria. In *Pt*, ConA and WGA bind to the haustoria in a similar pattern, although apparent changes in staining were observed as the haustoria mature, and the neck region fluoresced with either of the two lectin probes (Cantrill and
Chapter 2: Prediction and characterization of Pst effectors

Deverall, 1993). My observations with Pst were that WGA did not bind uniformly to haustoria, and this variability was independent of haustorial age. On the other hand, ConA bound more uniformly to haustoria of all ages, although the neck area did not fluoresce. These differences between haustoria of wheat rust fungi may reflect differences in wall chemistry, the relative preservation and availability of active binding sites after isolation, or even the resultant mix of ages and developmental stages of haustoria released by homogenization of infected wheat tissue.

2.4.2. Prediction and expression analysis of the Pst effector candidate genes

Effector proteins from fungal and oomycete parasites seem to travel through the eukaryotic endomembrane pathway for secretion (Ellis et al., 2009, O'Connell and Panstruga, 2006). This pathway involves exocytosis of Golgi-derived secretory vesicles (Korotkov et al., 2012). Most known fungal and oomycete effectors possess a canonical N-terminal secretion signal, which is sufficient for secretion (Dodds et al., 2004, Catanzariti et al., 2006, Panstruga and Dodds, 2009), and currently this is the most important and universal criterion to define an effector gene. From the 12,846 transcripts assembled here from the 454 haustorial transcriptome, ~19,000 ORFs were predicted from the 6 possible translation frames, and fewer than 1% of these were predicted to contain an SP for secretion. Although the SP is typical of proteins which are secreted from the cell, it can also target proteins into cellular membranes. It is important to distinguish between extracellular proteins and membrane-anchored proteins to avoid false predictions of secretion. So, further predictions of the presence of transmembrane domains helped to narrow the number of effector candidate genes to 437 (HSPs).

Effector genes in different fungal hemibiotrophic and biotrophic pathogens are usually expressed and often upregulated during infection (Stergiopoulos and de Wit, 2009), characteristics that are often used as criteria to define potential effectors. Initially I studied the expression of a subset of HSPs by RT-PCR, using cDNAs from different tissues, including germinated spores as a base state to compare against haustoria. Unexpectedly, a significant fraction of the HSPs (25%) were found to be equally expressed in germinated spores and haustoria, despite the fact that all HSPs were predicted from the haustorial transcriptome. An extended analysis of HSP expression was performed by using deep transcriptome sequencing (Illumina) of haustoria and germinated spores. The results of this analysis revealed four groups of expression: (i) exclusively expressed in haustoria, (ii) up-regulated in haustoria, (iii) non-differentially expressed and (iv) up-regulated in spores. The largest fraction of the HSP genes were classified in groups i and ii, supporting the idea that most of the effector proteins
synthesized in the haustorium play roles only during the biotrophic phase of the infection. However, similar to the results obtained with RT-PCR, approximately 28% of the HSP transcripts were present in germinated spores (expressed or up-regulated). Recently, evidence from proteomic data of germinating conidia of B. graminis suggested that effector proteins may be expressed very early during penetration, and that these might be already present in the germinating conidia before haustoria are formed (Pedersen et al., 2012). This classification of HSPs according to their patterns of expression is a valuable method of discrimination. For example, it can be used to inform the choice of candidates for functional analyses. It can also guide studies of the HSP promoters and associated mechanisms of temporal and differential regulation of expression, and perhaps could even be a tool to associate their expression with their genomic location or higher-level organization (Kamper et al., 2006). In Ustilago maydis for example, effector genes are clustered in the genome and most of these clusters are regulated together and induced in infected tissue (Kamper et al., 2006). The preliminary genome assembly of the Pst strain used in this study (Jackson et al., unpublished data), combined with the current work, offers an excellent opportunity to unravel the molecular strategy of Pst in deployment of its effector genes to cause disease.

2.4.3 Sequence analysis of the predicted Pst HSPs

To obtain a better understanding of the HSPs, they were compared with public databases extensively using BLAST searches to help identify them. Notably, 84% of the HSPs could not be annotated by homology-based inference as they do not have similarity to any known genes/protein, or because they are similar only to hypothetical unannotated proteins from recently sequenced fungal genomes. This lack of homology appears to be a feature of fungal effector sets predicted from other pathogens (Duplessis et al., 2011, Saunders et al., 2012, Pedersen et al., 2012). Only 9% of the genes could be annotated based on their similarity to known genes, and another 8% based on the presence of conserved protein domains. Most HSPs annotated by similarity were enzymes related to degradation or modification of fungal or plant cell walls, in particular enzymes related to chitin modification or degradation. Cell wall degrading enzymes are a hallmark of filamentous pathogen secretomes (Tian et al., 2009, Mueller et al., 2008, Raffaele et al., 2010), although there is a marked reduction in the number of genes encoding these kinds of enzymes in biotrophic fungal genomes relative to non-biotrophic (Duplessis et al., 2011, Spanu et al., 2010, Baxter et al., 2010). Nevertheless, biotrophic fungal pathogens must deal with the plant cell wall as a barrier to the initial stages of invasion and then locally to breach the host cell wall to form haustoria. Plant cell wall degrading enzymes have been detected and studied in the biotrophic pathogen U. viciae-fabae
(Deising et al., 1995). Interestingly, the regulation of these enzymes appears to be strictly controlled by the differentiation of infection structures rather than by substrate induction or catabolite repression, in contrast to necrotrophic fungal pathogens (Heiler et al., 1993, Deising et al., 1995).

HSPs related to oxidoreduction activity, ribonuclease activity and sequence-specific transcription factor activity were also detected. In Phytophthora infestans, an oomycete hemibiotrophic plant pathogen, proteins with oxidoreduction and ribonuclease activity are particularly enriched in its effector set (Raffaele et al., 2010), but their specific roles in infection are unknown. The ribonuclease-like domain is also enriched in the predicted effector set of B. graminis, and it was hypothesised that effectors with this domain could interact with host RNAs and modulate host immunity (Pedersen et al., 2012). Most of the domains found in HSPs were previously identified in effector sets from other biotrophic fungal pathogens (Duplessis et al., 2011, Saunders et al., 2012). These include; the CEFM domain, the Allergen V5/Tpx-1-related domain, and the thaumatin domain. Proteins containing a CEFM domain, an eight-cysteine-containing domain unique to the fungal kingdom, have been suggested to play a role promoting pathogenicity. These proteins are highly expressed during infection in B. graminis (Grell et al., 2003), M. oryzae (Xue et al., 2002) or Melampsora spp. (Joly et al., 2010), and M. oryzae mutants in genes encoding such proteins show reduced appressorial penetration (Xue et al., 2002). Observations in other pathosystems as well as in some symbiotic fungi suggest that proteins containing this motif could function as cell-surface receptors or signal transducers, or as adhesion molecules connecting the fungus and the host. The thaumatin domain is found in pathogenesis-related proteins with antifungal activity. Thaumatin-like proteins are induced in plant as a result of pathogen attack, elicitor treatments, environmental stress, and developmental signals (Datta et al., 1999), and some are associated with a protein kinase domain and were proposed to act as receptor-like kinases during defence responses (Wang et al., 1996). Thaumatin-like secreted proteins of rust fungi could perhaps alter such a plant signalling pathway, but further work must be done to clarify their role(s) during infection.

Although the search for conserved motifs in the predicted HSP protein sequences failed to identify conserved sequences, the search for conserved cysteine residues in fixed positions with the sequences allowed me to group a subset of HSPs. Similar analyses done on predicted effectors from genomic sequences of other rust fungi have shown that a high percentage of these genes are Cys-rich, and groupings according to their cysteine patterns were also possible (Hacquard et al., 2012, Duplessis et al., 2011). Furthermore, analyses of families of effector candidates that were under positive evolutionary selection (new advantageous genetic variants that increase the prevalence of
the beneficial allele) showed that the cysteine patterns were conserved despite low overall sequence identity (Hacquard et al., 2012). This suggests that the cysteine residues are conserved to allow significant changes in the remaining amino acids without altering the overall fold topology, leading to rapid diversification and the emergence of new virulences (Hacquard et al., 2012).

2.4.4. Functional characterization of Pst HSPs

Functional characterization of effector proteins is one of the most challenging areas in the study of biotrophic organisms. However, it is still the only way to differentiate real effector proteins from secreted proteins not related to pathogenesis. There are two major factors which limit functional analyses. First, as most of these proteins do not show homology to annotated proteins, it is not possible to predict biochemical functions that could be tested. Second, efficient transformation methods have not yet been developed for most of the strict biotrophic fungi including Pst, so tests have to be conducted in heterologous systems. However, such approaches lack context. One of the approaches used to characterize effector proteins has been to predict roles that could be necessary for a pathogen with a given strategy for pathogenicity, and then screen candidate effectors to identify those which can fulfil that role (Fabro et al., 2011, Badel et al., 2013). The fact that most effector genes are highly expressed in the haustorium and some are recognized by cytoplasmic R proteins inside the host cell (Dodds et al., 2004, Catanzariti et al., 2006), suggested that haustoria-forming pathogens deliver effector proteins into the host cytoplasm where they exert their roles in virulence. Thus, the use of the bacterial effector delivery system to study fungal effectors in a high-throughput manner has proven to be a useful alternative to inquire into the functions of effector proteins function (Yin et al., 2011b, Upadhyaya et al., 2014), especially in wheat rust fungi where transient gene expression assays are very difficult.

In this study, the P. fluorescens effector-delivery method (Thomas et al., 2009) was used to investigate potential recognition of HSPs in different wheat cultivars (R-Avr recognition assay). The assay was established and applied to only 5% of predicted effectors due to time constraints, and no conclusive results were obtained. The large number of effector candidates predicted for rust fungi from genomic and transcriptomic data (Garnica et al., 2013, Saunders et al., 2012, Duplessis et al., 2011), and the large proportion that are evidently deployed during the biotrophic phase of the infection (Garnica et al., 2013), suggests that each of these candidates would have similar possibilities of being recognized by a host. There are no directly identifiable features that indicate whether effector genes will be recognized by plant R genes. Thus, for this strategy to be effective,
larger numbers of HSP genes must be tested to increase the probability of finding a recognition event. Additionally, the *P. fluorescens* effector-delivery method assumes that the recognition event will occur in the host cytoplasm, as has been suggested for characterized effectors in other rust fungi (Catanzariti et al., 2006, Stergiopoulos and de Wit, 2009). However, some plant pathogens secrete effectors into locations other than the host cytoplasm, to facilitate colonization and suppression of host defences (Chisholm et al., 2006, Song et al., 2009, van Esse et al., 2008, Kamoun, 2006, Thomma et al., 2005). This possibility is still an open hypothesis for rust fungi and indeed seems likely. My study showed that HSP genes were expressed not only in haustoria but also in geminating spores, long before haustoria are produced, consistent with this idea. It cannot be ruled out that after synthesis in haustoria, effectors could be translocated to other fungal structures (i.e. infectious hyphae, newly forming spores) and secreted from there (this possibility is discussed further in Chapter 5). Thus, while the *P. fluorescens* effector-delivery method is useful to screen effectors that can be recognized intracellularly, it gives no information on those effectors which have a role outside of the cell. Additional factors such as proper protein folding and the relative level of fungal effector expression from the contrived bacterial system are important points that are yet unknown. However, the lack of alternative strategies for manipulating effector expression in planta, particularly genetic transformation of rust fungi, necessitates use of the bacterial system.

On the other hand, the *P. fluorescens* effector-delivery method was used to investigate the ability of HSPs to inhibit cell death induced by the *S. nodorum* toxin ToxA. Initial tests on *Pst* infected tissue suggested that the fungus had an inhibitory effect on ToxA-induced necrosis, leading me to test if this effect was a function of certain effector proteins. When I tested the inhibitory effects of a small set of HSPs using *P. fluorescens*, I found that three of them could delay and partially suppress the necrosis induced by ToxA. Although it is still unknown how exactly ToxA triggers cell death, and what the role of the susceptibility gene *Tsn1* is, the available evidence points to disruption of photosynthesis. ToxA is internalized into ToxA-sensitive but not -insensitive mesophyll cells. Once internalized, it localizes to chloroplasts (Manning and Ciuffetti, 2005), binds a photosystem II protein called plastocyanin (Tai et al., 2007), and its activity is light dependent (Manning and Ciuffetti, 2005). Thus, since *Pst* possess the capacity to partially suppress ToxA effect, one possibility could be that individual effectors translocated to the host cytoplasm inhibit the ToxA-induced cell death either because they block signal transduction pathways resulting in cell death, or enhance or protect photosynthesis so that it continues to provide nutrients to sustain fungal growth. Alternatively, *Pst* effectors could prevent ToxA from being perceived by the host, could block its internalization into
the chloroplast, or even could induce its degradation. The identification of interactors or targets of these HSPs in the plant cell could give some insights into the mechanism of these candidate effectors to inhibit ToxA-induced necrosis.

Suppression of plant immunity has emerged as a primary function of effectors, a concept that has been demonstrated for many TTSS effectors of plant pathogenic bacteria (Abramovitch et al., 2006, Jones and Dangl, 2006, Zhou and Chai, 2008, Hann et al., 2010), as well as for effectors of some fungal and oomycete plant pathogens. Several molecular strategies have been proposed for effector proteins from different organisms. Effectors may mask the pathogen from recognition as is the case for the LysM effector EcP6 in *Cladosporium fulvum*, which sequesters chitin oligomers originating from the fungal cell wall and therefore prevents PAMP-triggered immunity (de Jonge and Thomma, 2009). They can also inhibit the host defence proteins as is the case for *C. fulvum* Avr2, which bind the host cysteine proteases RCR3 and PIP2 to suppress immunity (van Esse et al., 2008). More recently, the Pep1 gene from *Ustilago maydis* which is essential for the establishment of biotrophy, was observed to inhibit the plant oxidative burst, acting as a potent suppressor of early plant defences (Hemetsberger et al., 2012). Shy1, a cytoplasmic salicylate hydroxylase also from *U. maydis*, is expressed during the biotrophic phase and was found to degrade salicylic acid possibly as a strategy to lower the amounts of this defence hormone (Rabe et al., 2013). On the other hand, AvrPto from the bacterium *Pseudomonas syringae* interferes with the flagellin-triggered signaling cascades mediated by the receptor kinase FLS2 (Xiang et al., 2008). The effector protein AvrPiz-t, which is secreted by pathogenic hyphae of the fungal rice pathogen *Magnaporthe oryzae*, suppresses BAX-induced programmed cell death in tobacco leaves (Li et al., 2009). In all these strategies, effectors target different proteins or signaling cascades to enhance pathogen fitness. The ToxA screening method developed in this work opens new possibilities to investigate novel roles of fungal effector proteins, and could guide further exploration on the function of the three effector candidates identified here.
Chapter 3
Transcriptome analysis of *Pst*

3.1. INTRODUCTION

The haustorium is a site of concerted host-pathogen interaction, so describing its functions is essential to understanding biotrophy. The presence of sugar transporters (Voegele et al., 2001) and putative amino acid transporters (Hahn and Mendgen, 1997, Hahn et al., 1997b) within the structure implies an important role in fungal nutrient uptake. However, other basic questions regarding the function of the haustorium have yet to be addressed. These include the identification of mechanisms for bidirectional transport (importation of nutrients and export of effectors and other molecules), the pathogen and host proteins specifically located at the haustorium–host cell interface, and the regulatory genes which specify haustorial identity.

The availability of genomic data for obligate biotrophs has increased dramatically in the last decade providing important insights into their infection strategies. For example, analysis of the genome sequences for the plant pathogens *Puccinia graminis* f.sp. *tritici* (*Pgt*), *Melampsora larici-populina* (*Mlp*) (Duplessis et al., 2011) and *Blumeria graminis* (Spanu et al., 2010) have revealed that they share evolutionary features of adaption to the extreme parasitic lifestyle, such as the loss of nitrate and sulfate assimilation pathways. The genome of *Ustilago maydis*, a biotrophic pathogen of maize,
revealed that this fungus lacks cell-wall-degrading enzymes and possesses clusters of genes encoding small secreted proteins which have a decisive role in the infection process (Kamper et al., 2006). The genome of the obligate biotroph, *Blumeria graminis* f.sp. *hordei*, displays massive retrotransposon proliferation, absence of genes for secondary metabolic enzymes, and a reduction in genes encoding specific subfamilies of transporters, hallmarks of its exclusively biotrophic life-style (Spanu et al., 2010). Although a draft genome sequence of *Pst* has been published (Cantu et al., 2011), this is a rather limited study that does not provide similar insights into gene content and utilisation.

In addition to genomics, large-scale transcriptomic analyses are an effective way to investigate the gene expression program of a particular tissue or life stage of organisms (Wang et al., 2009c). Transcriptomics approaches based on next-generation sequencing technologies (known as RNA-seq) is an emerging tool to study unculturable biotrophic fungal pathogens. Examples of this are the recent transcriptomic studies on the powdery mildew fungus *Golovinomyces orontii* (Wessling et al., 2012) and the coffee rust fungus *Hemileia vastatrix* (Fernandez et al., 2012), which aimed to use transcriptomics to better understand biotrophy. In the first example, the transcriptome of enriched haustorial complexes revealed that gene products with roles in protein turnover, detoxification of reactive oxygen species and fungal pathogenesis are abundant in the haustorium, and also allowed the prediction of secreted proteins which comprise the effector candidates. Interestingly, this study showed that nutrient transporter transcripts were not abundant in the haustorial transcriptome, suggesting different regulation of nutrient acquisition in powdery mildew relative to rust fungi. In the second example, the transcriptome of *H. vastatrix* was analyzed during a compatible interaction with coffee plants to obtain an exhaustive repertoire of the genes expressed during infection and to identify potential effector genes. This study provided the first genomic resource for this agriculturally important plant pathogen. With further genomic datasets for biotrophic plant pathogens emerging, transcriptomic data has also proven useful for gene prediction and annotation (W. Jackson, The Australian National University, personal communication, (Spanu et al., 2010, Duplessis et al., 2011)).

Despite the agricultural importance of *Pst*, the available genomic resources are limited. Expressed sequence tag (EST) libraries have been generated for *Pst* (Ling et al., 2007, Ma et al., 2009, Zhang et al., 2008, Xu et al., 2011) and other *Puccinia* species (Thara et al., 2003, Hu et al., 2007, Broecker et al., 2006), which have been useful for stage-specific expression analysis and gene prediction in genomes. However, the resolution power of these studies, firstly for completeness, and secondly for comparisons of transcript abundance between cell types, is limited. A pressing need for a better
understanding of the virulence strategies of this fast evolving pathogen, together with the availability of high-throughput sequencing technologies, motivated the development of this part of the project. RNA-seq technologies based on the 454 and Illumina platforms enables the sequencing of whole transcriptomes and the accurate measuring of gene expression in virtually any biological sample (Maxmen, 2013). Therefore, taking advantage of these technologies, here I have combined haustoria purification methods (Hahn and Mendgen, 1992) with 454 and Illumina sequencing to broadly analyze gene expression in the germinated uredospores (henceforth called “germinated spores”) and haustorial stages of *Pst*. The transcriptome sequencing of both pathogenic stages through Illumina and 454 platforms allowed the generation of a set of reference transcripts which were subsequently used as templates to perform a digital expression analysis. Automatic and manual annotation of the assembled transcripts was performed to determine their identity and to classify them into metabolic pathways or functional categories. Finally, the expression levels in haustoria and germinated spores were integrated into the analysis to put all the information into biological context. The global comparison of gene expression in haustoria and germinated spores revealed fundamental metabolic differences between the two developmental stages, and gave new insights into *Pst* pathogenesis strategies.

3.2. MATERIALS AND METHODS

3.2.1 Semi-purification of *Pst* haustoria using Percoll gradients

For Illumina RNA-seq expression profiling, haustoria were purified by a protocol that I developed using density gradients in a colloidal silica media called Percoll (described in Chapter 4). Briefly, 20 g of infected wheat leaves were surfaced sterilised as in section 2.2.2.1 from chapter 2, blended in 250 ml of 1x isolation buffer [0.2 M sucrose, 20 mM MOPS pH 7.2] (1x IB), and passed consecutively through 100 µm and 20 µm meshes to remove cell debris. The filtrate was centrifuged at 1080 g for 15 min at 4°C and the resulting pellets were resuspended in 80 ml of 1x IB containing 30% Percoll v/v. The suspension was centrifuged at 25,000 g for 30 min at 4°C. The first 10 ml of each tube was recovered, diluted 10 times with 1x IB and centrifuged at 1080 g for 15 min at 4°C. The pellets were resuspended in 20 ml of 1x IB containing 25% Percoll and centrifuged at 25,000 g for 30 min at 4°C. The first 10 ml of each tube was recovered, diluted 10 times with 1x IB and centrifuged at 1080 g for 15 min at 4°C. The final pellets were frozen in liquid nitrogen and stored at -80°C prior to RNA isolation.
3.2.2. RNA isolation and sequencing from Pst haustoria and germinated spores

Frozen haustoria pellets derived from a single isolation were processed individually, ground with an electrical tissue homogenizer, and total RNA was isolated using the QIAGEN (Doncaster Australia) Plant RNeasy kit following the manufacturer’s instructions. The resulting RNA was checked for concentration and possible contamination on a NanoDrop ND-1000 UV-Vis Spectrophotometer and subsequently treated with DNase I (New England Biolabs, Beverly, MA) following the manufacturer’s instructions. Samples were purified using the QIAGEN Plant RNeasy kit following the cleanup protocol, and RNA was eluted from columns in 50 µl of RNase-free water. Three biological replicates were processed.

To prepare germinated spores, ~100 mg of freshly collected spores were sprayed evenly onto MilliQ water in a glass petri dish (15 mm diameter, 8 cm high) and incubated in darkness for 15 h at 9°C. Germination was verified by microscopic examination, and germinated spores were recovered by filtering the water containing the spores through an 11 µm nylon mesh. Retained material was dried as much as possible with tissue paper and snap frozen in liquid nitrogen. Three independent samples were ground in liquid nitrogen, and total RNA isolation, DNAse I treatment and RNA clean-up were performed as for haustoria tissue. The concentration and integrity of all haustoria and germinated spores RNA samples were verified on the Agilent 2100 bioanalyzer, using the RNA 6000 nano and pico kits.

Haustoria and germinated spores RNA samples were sent to The Ramaciotti Centre (University of New South Wales, Sydney, Australia) for sequencing with the Illumina HiSeq 2000 instrument, 100-bp paired end reads. Approximately 10 µg of total RNA per biological sample was processed with the TruSeq RNA Sample Preparation Kit v2. High sensitive DNA assay was done on the Qubit® 2.0 Fluorometer (Life Technologies) to quantify the generated cDNA libraries and additional quality controls were applied by running the libraries on the Agilent 2100 Bioanalyzer using a DNA 1000 chip kit to verify that they are the expected size and there are no adapter dimer bands.

3.2.3. Reference transcriptome assembly and BLAST2GO and BLASTx annotation

CLC Genomic Workbench 4.0 software (http://www.clcbio.com/) was used for de novo assembly of combined 454 data from haustoria and germinated spores (accession number SRR579539-40) previously described in Chapter 2. The parameters used for de novo assembly were similarity 0.97;
length fraction 0.5; insertion cost 3; deletion cost 3; mismatch cost 2; minimum contig length 200 nt. A total of 12,282 transcripts were assembled. This dataset comprises the reference transcripts set referred to in the remainder of this chapter. For the gene ontology (GO) classification the reference transcripts set was analyzed using BLAST2GO (B2G) 2.5.1 (Conesa et al., 2005). Parameters were set to maximum e-value < 10^{-25}, maximum number of alignments to report = 20 and highest scoring pair length = 33 amino acids. Additionally, BLASTn and BLASTx screening against NCBI non-redundant nucleotide and protein databases were used for manual analysis. Only the transcripts that belonged to one of the functional categories were analyzed manually to verify if more than one transcript represented the same gene, in which case it was annotated in Tables S4-S7. For the remainder of the transcripts, which comprise sequences without homology to annotated genes/proteins, this was not possible because different transcripts might have the same top hit without meaning that they were derived from the same gene.

3.2.4. Digital expression analysis

For digital gene expression analysis, Illumina RNAseq reads from haustoria and germinated spore mRNA samples were mapped against the reference transcripts set using the RNA-seq tool of CLC genomics 5.0. Illumina reads were first processed by removing adapter sequences and reads of low quality (trim quality score 0.05, maximum nucleotide ambiguities = 2 and minimum number of nucleotides in reads = 35), allowing up to two base mismatches. The parameters used for the mapping were minimum length fraction 0.8, minimum similarity fraction 0.97, maximum number of hits for a read 10, inclusion of broken pairs in the counting scheme, and computation of normalized counts in RPKM (expressed in reads per kilobase of exon model per million mapped reads (Mortazavi et al., 2008)). Reads that mapped to multiple sites were assigned randomly and proportionally to one of the mapped sites, according to the number of unique matches that the genes to which it matches have, normalized by the transcript length. Three biological replicates for each pathogenic stage were used in the differential gene expression analysis. In order to determine inner dataset dynamics and the reproducibility among biological replicates, Pearson correlation coefficients were calculated for haustoria and spores replicates. Principal component analysis (PCA) was also done to evaluate the similarities and differences between the 6 transcriptomic profiles (3 from haustoria and 3 from spores). Both, Pearson coefficients and PCA analyses were done with CLC genomics. The differential gene expression between haustoria and germinated spores samples was evaluated using Baggerley’s test (Baggerly et al., 2003) by treating the same types of sample as one group. Those genes with a false discovery rate (FDR) corrected \( P \)-value of less than 0.05, a fold induction of >2 and
a difference in the RPKM average value of at least 20 were considered to be significant. B2G was used for GO functional enrichment analysis of the transcripts that were differentially expressed in both spores and haustoria, by performing Fisher’s exact test with false discovery rate (FDR) correction to obtain an adjusted \( P \)-value. Illumina datasets were deposited in the NCBI Sequence Read Archive under the accession numbers SRR579533-38.

3.3. RESULTS AND DISCUSSION

Previously I used 454 sequencing of isolated haustoria and germinated spores to facilitate the de novo assembly of transcripts which were subsequently analyzed to predict effector candidates, as described in Chapter 2. However, the large amount of data generated, the availability of data-analysis software for non-specialists, and the development of a much more efficient technique to isolate haustoria (described in Chapter 4), allowed me to explore the content of the transcriptomic data further. Extensive annotation of this dataset, complemented with Illumina RNA sequencing to quantitate levels of gene expression, provide great insight into the pathogenic strategies behind different gene expression programs used by this organism.

3.3.1 Assembly of a reference transcriptome for gene expression analysis

The comparison of transcriptomes to determine biologically meaningful differences relies on three main factors: (i) The availability of a genomic or transcriptomic reference, (ii) the data to assess the levels of expression of each particular gene/transcript and (iii) information about the identity of the studied genes/transcripts. A comprehensive transcriptome-scale analysis of \( Pst \) using publicly available genomic resources was not possible at the time this study started as there were no reference genome or commercial array platforms for expression profiling of this pathogen. The only public resources for \( Pst \) were small EST collections from different tissues (Ling et al., 2007, Zhang et al., 2008, Yin et al., 2011a), all derived from different laboratories and from different strains, making it very difficult to associate trends in gene expression with specific biological functions. Recent breakthroughs in next generation sequencing technology and data analysis suggested that it would be possible to generate a reference transcriptome in the absence of a reference genome (Grabherr et al., 2011, Mutasa-Gottgens et al., 2012, Sanchez et al., 2011), and then to use this reference transcriptome to perform comparative studies by methods such as digital gene expression profiling (Mutasa-Gottgens et al., 2012). Thus, I pooled the 454 raw reads previously generated from isolated haustoria with those from germinated spores and assembled them de novo to create a reference
transcriptome of 12,282 transcripts (Table S3, Transcripts reference set and Data S5), which represent the cDNAs of genes expressed during early and advanced stages of the infection.

My next objective was to assess the transcriptional levels of the assembled transcripts in isolated haustoria and germinated spores. Some studies have shown that 454 pyrosequencing datasets can be used for quantitative estimation of gene expression through the relative abundance index, which is based on number of 454 sequences assembled into a given contig normalized for the length of that contig (Vega-Arreguin et al., 2009, Fernandez et al., 2012, Torres et al., 2008). However, although the 454 platform produces the longest reads among the most widely used NGS platforms, there is more recent evidence that quantifying transcript abundance is best served by generating shorter but more abundant reads (Siebert et al., 2011, Zagordi et al., 2012). Thus, I chose the Illumina HiSeq 2000 sequencing platform to carry out a digital expression analysis for quantitation of transcript abundance. This platform reaches a much higher coverage level per run than 454, increasing the quantitative power and allowing statistically robust quantitative comparisons.

3.3.2. Isolation of Pst haustoria using Percoll gradients, and RNA extraction from haustoria and germinated spores

To sequence haustoria and germinated spores on the Illumina platform, new sample preparations were necessary to generate the three biological replicates from each pathogenic stage. Although I previously purified haustoria using an affinity chromatography method (Hahn and Mendgen, 1992), this technique generates haustoria with low yield and purity so was not pursued here. Instead, I developed a new method using Percoll gradients to purify haustoria based on their density, which significantly improved the yield of the procedure and is described in detail in Chapter 4.

Haemocytometer counts of isolated haustoria by Percoll density gradients showed that 5x10^6 to 6.5x10^6 haustoria could be isolated from a single infection (~15-20 g of infected tissue), from which >5 µg of total RNA per sample could be purified, which was sufficient for the preparation of Illumina sequencing libraries. Total RNA from isolated haustoria and germinated spores was extracted concurrently using the same column purification method. Preliminary assessments of RNA concentration and possible contamination were done on a Nanodrop spectrophotometer, however further evaluations on the sizing, quantity and quality of the RNA were done on the Agilent 2100 Bioanalyzer, using RNA 6000 nano and pico kits (Kuschel, 2000) (Figure 3.1).
Chapter 3: Transcriptome analysis of Pst

Figure 3.1 Bioanalyzer electropherograms of total RNA. “Live” plots of nucleic acid fluorescence against migration time of total RNA preparations from Pst tissues. During chip preparation, the micro-channels are filled with a sieving polymer and a fluorescent dye. Dye molecules intercalate into RNA strands and these complexes are detected by laser-induced fluorescence. The data are converted into electropherograms. A standard curve for migration time versus fragments size is plotted with reference to molecular weight standards. The size and concentration of the nucleic acids in the sample are calculated from the migration times measured for each fragment in the sample and their fluorescence intensities. A. Germinated spores, B. Percoll-isolated haustoria and C. Percoll-isolated uninfected tissue. Information derived from the electropherograms such as RNA concentration, rRNA ratio and RIN is shown at the bottom of each image.

The RIN is a Bioanalyser tool designed to assist with estimating the integrity of total RNA samples. It takes into account a selection of features that give information about the RNA integrity, for example the areas under the curve of the 18S and 28S peaks compared to the total area under the curve, the height of the 28S peak, and the height of the lower marker added to RNA samples, amongst others (Schroeder et al., 2006). Total RNA from all germinated spore preparations were of very high quality, as indicated by the electropherogram pattern (Figure 3.1A) and RIN which was always very close to 10 (highest quality showing the least degradation) (Mueller et al., 2004). However, total RNA from isolated haustoria consistently showed a low quality pattern (Figure 3.1B) as the 18S and 28S peaks were not clearly defined and RIN was <7, which could result in the failure to generate cDNA libraries. To investigate this further, I performed several further haustorial extractions, and included spores and infected tissue samples as positive controls for the extraction process to help identify when and how RNA was apparently being degraded. All of the extractions from the positive controls produced very high quality RNA, but low quality RNA from isolated haustoria. Haustoria isolated on Percoll gradients are mainly composed of haustoria and chloroplasts at a ratio of ~1:60 (Chapter 4). In this case, it is possible that the large number of contaminating chloroplasts could be responsible for the low quality electropherogram. To investigate this, I processed uninfected wheat tissue using the same method to isolate haustoria from infected tissue, extracted the RNA from the comparable
Percoll fraction containing only chloroplasts, and analyzed the quality of the RNA on the Bioanalyzer (Figure 3.1C). The pattern obtained partially matched that of the haustoria sample, suggesting that the high chloroplast content does contribute to the haustoria RNA electropherogram pattern. Similar results have been obtained from semi-pure haustoria samples of other wheat rust pathogens (N. Upadhyaya, CSIRO Plant industry, personal communication), indicating that this could be an effect of the isolation method. Alternative explanations such as an intrinsic characteristic of the particular sample, or sample degradation, cannot be ruled out. Nevertheless, the multiple experimental attempts and controls performed to find out the origin of this pattern suggested that the former option is improbable (discussed later). In the absence of other options, the haustoria RNA samples (three biological repeats) generated in this way were sent for RNA-seq Illumina HiSeq 2000 sequencing to The Ramaciotti Centre (University of New South Wales, Sydney, Australia), along with RNA isolated from germinated spores (also three biological replicates).

### 3.3.3 Illumina sequencing

More than 500 million paired-end reads (average insert 146 bp) were obtained from Illumina Hiseq 2000 sequencing of mRNAs from germinated spores and purified haustoria (Table 3.1), which were subjected to strict quality control procedures. Every NGS experiment consists of a series of steps that contribute to the overall quality of the data, and different metrics have been developed to measure the accuracy of each step. Base calling accuracy is one of the most critical data quality indicators and is commonly measured by the Phred quality score (Q Score), which estimates the probability that a base is called incorrectly (Ewing et al., 1998). Q scores are defined as a property that is logarithmically related to the base calling error probability (P). The quality score for each base ranges from -5 to 40 and is defined as: $Q = -10 \log_{10} P$ (Ewing et al., 1998). Higher quality scores indicate a smaller probability that a base is called incorrectly. For example, if Phred assigns a Q score of 30 (Q30) to a base, this is equivalent to the probability of an incorrect base call 1 in 1000 times (i.e. probability of a correct base is 99.9%). A lower base call accuracy of 99% (Q20) will have an incorrect base call probability of 1 in 100. When sequencing quality reaches Q30, virtually all the reads will be perfect, having zero errors or ambiguities, so this score is considered a benchmark for quality in NGS (Illumina, 2011).

To determine the distribution of quality scores in my data, Illumina quality scores for each library were transformed to the Phred scoring scale and plotted according to the base position along the
read. Figure 3.2 shows examples of the quality scores graphs for one of the libraries for isolated haustoria, and one for germinated spores.

Figure 3.2 Average Phred quality score per base for haustoria and spores libraries before quality control.

All of the libraries looked very similar, with very high average Q scores in almost every base position, consistent with high quality data. To filter the data according to Phred scores, terminal spans of bases with Phred scores less than 20 were clipped using the CLC genomics quality control tool. I used the standard arbitrary criterion of at least 90% of bases in each read exhibiting a Phred score of 21 or higher for the read to be retained for further analyses. Also, ambiguity trimming (e.g. stretches of lateral Ns) and adapter trimming were used to clean the data. The total number of reads per library before and after quality control is shown in Table 3.1.

3.3.4. Digital expression analysis

To compare gene expression between haustoria and spores, raw Illumina reads comprising independent biological replicates from isolated haustoria and germinated spores were independently mapped against the transcriptome reference gene set, using the RNA-seq tools from CLC genomics (Table 3.1). While the total number of reads from each sample and pathogenic stage varied, especially for isolated haustoria, approximately 60% to 75% of the reads per library could be mapped back to the assembled transcriptome.
Table 3.1. Illumina sequencing data before and after mapping against the unigene reference. Three biological replicates were sequenced with Illumina for isolated haustoria (H) and germinated spores (S). The table shows the millions of reads obtained per replicate and the percentage of reads mapping against the reference unigenes assembled from 454 haustoria and germinated spore data.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Original reads (Mill)</th>
<th>After trimming (Mill)</th>
<th>Mapping in pairs (Mill)</th>
<th>Mapping in broken pairs (Mill)</th>
<th>% Mapping</th>
<th>Not mapping (Mill)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>38.5</td>
<td>37.5</td>
<td>20.4</td>
<td>1.8</td>
<td>60</td>
<td>14.3</td>
</tr>
<tr>
<td>H2</td>
<td>70.3</td>
<td>68.3</td>
<td>41.7</td>
<td>4.2</td>
<td>67.2</td>
<td>22.3</td>
</tr>
<tr>
<td>H3</td>
<td>53</td>
<td>51.9</td>
<td>30</td>
<td>2</td>
<td>61.6</td>
<td>18.7</td>
</tr>
<tr>
<td>S1</td>
<td>110.2</td>
<td>108.1</td>
<td>73.7</td>
<td>7.1</td>
<td>74.7</td>
<td>25.3</td>
</tr>
<tr>
<td>S2</td>
<td>120.5</td>
<td>118.1</td>
<td>79.6</td>
<td>8</td>
<td>74.1</td>
<td>28.1</td>
</tr>
<tr>
<td>S3</td>
<td>115.6</td>
<td>113.4</td>
<td>77.9</td>
<td>7.4</td>
<td>75.2</td>
<td>25.9</td>
</tr>
</tbody>
</table>

The behaviour of the biological replicates as sample groups were visualised by principal component analysis (PCA). PCA is a mathematical analysis that identifies and quantifies the directions of variability in the data. It allows the identification of groups of variables (components) that are interrelated via phenomena that cannot be directly observed, and that explain the maximum amount of mutual correlation (Burstyn, 2004). The plot in Figure 3.3 is based on a two-group experiment (haustoria and germinated spores) and shows the projection of the samples onto the two-dimensional space spanned by the first and second principal components. The first component is the direction through the data that explains the most variability in the data. The second and subsequent component must be orthogonal (at right angles) to the previous component and describes the maximum amount of the remaining variability. Each sample has a score for each component and is represented by the coloured dots in figure 3.3.
Figure 3.3. Principal Component Analysis of RNAseq replicate samples. Each dot represents one biological replicate and is coloured according to the pathogenic stage; red for isolated haustoria, green for germinated spores.

In this scatter plot, samples belonging to the same pathogenic stage group clustered together, exhibiting less variability than between samples from different groups. The first component explains 62% of the variance while the second component explains the 34%. Thus, at an overall level the distributions of the sets of expression values in biological replicates were more similar and clustered in groups, showing that the RNA-seq-derived transcriptome profiles from haustoria and spores are characteristic to different pathogenic stages (Figure 3.3). The PCA plot is thus helpful in identifying outlying samples and samples that have been wrongly assigned to a group. Pearson’s correlation coefficients were also calculated between replicates to assess the reproducibility of the sequencing. Correlation coefficients of 0.84–0.88 were recorded for haustoria biological replicates, and 0.88–0.93 for germinated spores biological replicates, indicating good reproducibility of the sampling procedure.

Sequence read counts were adjusted to reads per kilobase per million mapped reads (RPKM, (Mortazavi et al., 2008)) in CLC genomics, to enable the comparison of transcript levels assigned to each gene between samples. RPKMs were statistically assessed using Baggerley’s test (Baggerly et al., 2003) (FDR-adjusted p-value < 0.05) for each gene to determine differential expression between haustoria and germinated spores. A total of 4,346 transcripts were differentially expressed, revealing a clear difference between the transcriptional programs of germinated spores and haustoria (Table S3 in Garnica et al., 2013 and Figure 3.4).
Figure 3.4. Differential gene expression in germinated spores and haustoria of Pst. Venn diagram of the reference transcripts set, showing the number of transcripts that did not show differential expression between the two pathogenic stages, and those that had statistically significant changes in expression between germinated spores and haustoria.

At this point, with the availability of expression values for each transcript and before annotating the transcripts further, I did an additional quality control on the RNA-seq data. Since it was not possible to identify the reason for the low RIN values for haustorial RNA, I investigated if on the derived sequenced reads showed a similar pattern of distribution along the gene as high RIN value samples. A bias of read coverage to the 3’ end of genes has been associated with degraded RNA starting material (Sun et al., 2013, Khrameeva and Gelfand, 2012). Recently, software has been developed to study RNA-seq metrics for quality control, including graphic representations of coverage over the gene body (DeLuca et al., 2012, Wang et al., 2012). However, knowledge of the coordinates of all the gene bodies (i.e. start and stop codon positions) within the reference is a prerequisite, information that is out of the scope of this thesis. Thus, a simpler approach was undertaken by randomly choosing transcripts with very similar RPKM expression values in germinated spores and haustoria, and analysing the distribution of the mapped reads (Figure 3.5).
Figure 3.5. Mapping of Illumina reads from haustoria and germinated spores against transcripts from the reference set. The distribution of sequence reads along transcripts was visualized using the graphical tool of CLC genomics. Five transcripts with very similar expression values (RPKM, indicated in each case) in haustoria and germinated spores were chosen, and the reads that mapped against these sequences are represented in aggregate by the pink areas. The green arrows represent the size and orientation of the most probable open reading frames (determined using BLAST searches). All of the images are presented at the same scale.

In all the cases the coverage appeared to be biased to the 3’ end of the transcripts in haustoria and germinated spores, although this effect was slightly more pronounced in the haustoria data. Poly-A cDNA synthesis methods are prone to 3’ bias, and this is represented by corresponding depth of coverage in NGS experiments (Levin et al., 2010, Gibbons et al., 2009). (Tariq et al., 2011, Khrameeva and Gelfand, 2012). Thus, the occurrence of this phenomenon in my haustoria and germinated spores datasets is to a certain extent due to a systematic bias in the sequencing method. Additionally, although quantification was not possible, the absence of major differences between
the coverage from haustoria and germinated spores reads suggests that the low RIN of haustorial RNA did not have a profound effect on the uniformity of read distribution along the transcripts.

### 3.3.5. Functional classification of the transcriptome sequences

The transcripts in the reference set were categorised into functional classes using BLAST2GO to identify genes that encode proteins with known roles in cellular processes. Of this set, 4,485 transcripts could be annotated unambiguously. Using the list of genes previously found to be differentially expressed, Fisher’s exact test was applied to find functional categories over-represented in either developmental stage. The major functional categories are shown in Figures 3.6 and 3.7. Broadly speaking, processes up-regulated in germinated spores were representative of cell proliferation, such as cell cycle control and DNA and cell wall metabolism, whereas haustoria were more engaged in energy production and biosynthetic processes. Figure 3.6 refers to transcripts overexpressed in each pathogenic stage in the selected GO categories, whereas Figure 3.7 includes all transcripts classified under the selected GO categories and relative levels of expression were represented as colours. This is important because processes included within the same ontogenic category can have fundamentally opposed activities. The data are discussed in more detail with reference to fungal development and metabolism in the subsections below.
Chapter 3: Transcriptome analysis of Pst

Figure 3.6. Comparative ontological analysis of transcripts with statistically significant changes in expression between haustoria and germinated spores. Of the original transcript set, 30.2% (601) of the 1,989 haustorial-enriched genes and 47% (1,109) of the 2,357 genes up-regulated in germinated spores could be annotated with B2G. GO terms related to relevant biological processes are shown on the Y-axis. Percentages of genes differentially expressed in each pathogenic stage belonging to the nominated categories are shown on the X-axis.

Figure 3.7. Heatmaps representing relative levels of gene expression classified by metabolic categories. Relevant categories were selected from the B2G analysis. Genes belonging to these categories were listed and organized according to transcript expression values (RPKMs) in haustoria and spores. The colours are based on haustorial expression relative to the overall expression in both pathogenic stages; red indicates high expression in haustoria; green, expression in germinated spores. Specific subcategories are shown by the black arrows indicating the presence of specific transcripts discussed in the text.

3.3.5.1. Cell proliferation

3.3.5.1.1. DNA replication and cell cycle

The two sampled tissues represent very different developmental stages, with germinated spores involved in growth and division, whereas the haustorial stage is terminally differentiated. Generally,
transcripts classified into the categories of cell cycle, DNA replication and cell division were more highly represented in germinated spores (Figure 3.7). Genes involved in DNA replication and cell division included DNA polymerase subunits and replication factors. Cell division genes were clearly enriched in germinated spores, consistent with the idea that haustoria are non-dividing fully differentiated cells. For example, cyclins (proteins that regulate cyclin dependent kinases (CDKs) to control cell division), cyclin-dependent kinases, cohesins (protein complexes that regulate the separation of sister chromatids during cell division), septins (GTP binding proteins that provide structural support during cell division and compartmentalize parts of the cell) and genes associated with the control of mitotic phase progression were all up-regulated in germinated spores (Figure 3.7 and Table S6, HeatMap). Although no conclusions can be drawn from the expression of these genes about their function in vivo because both positive and negative roles in control of cell cycle are possible, the observations suggest that active cell cycle regulation occurs in germinated spores. One of the most interesting transcripts found to be up-regulated in germinated spores encodes a protein similar to the cyclin-dependent kinase Cdk5 (PST79_11595). The cyclin Pcl12 represents a polarity- and virulence-specific regulator in Ustilago maydis, complexing with the cyclin-dependent kinase Cdk5 for sustained polar growth (Flor-Parra et al., 2007). Although I found no Pcl12 homologs in my data (nor in the Pst 130 draft genome (Cantu et al., 2011)), Pst encodes other cyclin-like proteins that might regulate Cdk5 function. Homologs of U. maydis cyclin genes B1 and B2 were also found to be highly expressed in Pst germinated spores, with the latter one of the most strongly up-regulated genes of this category. The U. maydis genes are important in cell cycle regulation (Garcia-Muse et al., 2004), with mutants arrested after S phase, while overexpression of cyclin B2 generates cells with anomalous DNA content and premature entry into mitosis. In M. oryzae, degradation of the homologous cyclin B proteins CYC1 and CYC2 is necessary for mitosis exit and pathogenesis (Saunders et al., 2010). The connections between cell cycle, morphogenesis and virulence appear to be important for this pathogen to colonise plant tissue, and this may also be the case for rust pathogens.

Four transcripts with similarity to septin genes (PST79_4650, PST79_4652, PST79_4126, PST79_582) were found in my data, three of which were up-regulated in germinated spores. Septins constitute a cytoskeletal structure that is conserved in eukaryotes (Cid et al., 2001). In Saccharomyces cerevisiae, septins assemble as a ring that marks the cytokinetic plane throughout the budding cycle and this structure participates in different aspects of morphogenesis, such as cell polarity, localization of chitin synthesis, and the spatial regulation of septation (Cid et al., 2001). Recently Dagdas et al. (2012) showed that septin proteins in M. oryzae assemble in a ring-shaped structure at the base of
the appressorium together with F-actin and other cytoskeletal components to provide rigidity and negative membrane curvature for protrusion of the penetration peg into the host (Dagdas et al., 2012). Although Pst can often form appressoria over stomata prior to leaf penetration, this is not strictly necessary (Broers and Lopez-Atilano, 1996, Wang et al., 2011a, Zhang et al., 2008). Thus the role of septins in leaf penetration by Pst remains to be established.

3.3.5.1.2. Cell wall modification enzymes

Cell walls are complex polysaccharide structures composed mainly of cellulose, hemicellulose and pectin in plants (Carpita and Gibeaut, 1993), and chitin, glucans and other minor components in fungi (Bowman and Free, 2006). Glycoside hydrolases (GH), polysaccharide lyases, and esterases are enzymes that allow formation, remodelling or degradation of cell walls and play a fundamental role in plant-fungal pathogen systems (Soanes et al., 2007). The battery of carbohydrate-modifying enzymes derived from plant pathogenic fungi can vary enormously depending on the ecological niches occupied by the different species. Recent genomic studies on plant fungal pathogens have shown that the number of GH genes is correlated with the type of host interaction (Ipcho et al., 2012, Dean et al., 2005, Duplessis et al., 2011, Kamper et al., 2006, Spanu et al., 2010). Biotrophic fungi appear to possess a reduced number of GH enzymes, consistent with the need to minimize host cell wall damage to avoid triggering plant immunity (Spanu et al., 2010, Duplessis et al., 2011). I found a total of 53 transcripts with similarity to GHs belonging to 18 different families that were expressed in Pst (Table S3, General metabolism). The most abundant families were GH5, GH18 and GH47, which agrees with recent findings for Pgt and Mlp, where the GH5 and GH47 families were expanded (Duplessis et al., 2011). Thirteen Pst GHs were similar to enzymes implicated in plant cell wall degradation and interestingly, of these, two α-galactosidases (PST79_3645, PST79_3417) and three β-mannosidases (PST79_2043, PST79_8128, PST79_4682) were up-regulated in haustoria. The higher expression of mannosidases at advanced stages of infection has been reported in other rust fungi (Duplessis et al., 2011). Moreover, two glucanases from family 16 (PST79_11351, PST79_11352) were massively up-regulated in germinated spores, which may have implications for host invasion. In barley, alpha-galactosidase activity is essential during leaf development by loosening cell walls and in cell wall expansion (Chrost et al., 2007). In Arabidopsis PEN gene products are known to be involved in the antimicrobial delivery system executing apoplastic immune responses. PEN1 encodes a plasma membrane syntaxin, involved in the timely formation of localized thickenings on the interior face of the plant cell wall at sites of attempted fungal entry (Assaad et al., 2004). PEN2 encodes an atypical myrosinase that hydrolyzes 4-methoxyxnidol-3-
ylmethyglucosinolates for antifungal responses (Lipka et al., 2005), while PEN3 was shown to encode an ATP binding cassette (ABC) transporter likely involved in exporting PEN2-generated toxic compounds to sites of attempted fungal invasion (Stein et al., 2006). Loss of PEN1 compromises penetration resistance against the non-adapted powdery mildew as *B. graminis* f.sp. *hordei* and *Erysiphe pisi*, and loss of PEN3 is associated with enhanced susceptibility to the same pathogens (Stein et al., 2006, Lipka et al., 2005). It has been demonstrated that PEN1, PEN3 and other extracellular defense components are involved in the formation of both papillae (which contains callose) and haustorial encasements (Meyer et al., 2009). It possible that similar defense mechanisms against haustoria-forming pathogens exist in cereals, therefore it is imperative for the pathogen to overcome this defense in a compatible interaction. The higher expression of cell wall degrading enzymes in *Pst* may suggest that this is a fungal strategy to impede reinforcement of the host cell wall at the site of haustorial interaction. Additionally, eight *Pst* transcripts encode proteins similar to β-1,3-glucanases, four of which were up-regulated in germinated spores (PST79_5045, PST79_2573, PST79_9378, PST79_8761) and one in haustoria (PST79_5049). These could potentially play roles either in fungal cell wall remodelling or degradation of plant callose. The remaining GHs were similar to enzymes required for glucan synthesis and fungal cell wall remodeling (almost all up-regulated in germinated spores), posttranscriptional modification of cell wall proteins, chitinases (almost all up-regulated in haustoria) and glycogen breakdown.

Related to cell wall modification, I also found genes with strong similarity to chitin synthases (mostly up-regulated in germinated spores) and carbohydrate deacetylases (including chitin deacetylases), which were massively up-regulated in germinated spores. Chitin deacetylases are of particular interest because they convert chitin into the less rigid chitosan, potentially avoiding recognition by plant chitin receptors (Gueddari et al., 2002). The genomes of *Pgt* and *Mlp* (Duplessis et al., 2011) as well as the secretome of the symbiont *Laccaria bicolor* (Vincent et al., 2012a), were also found to be enriched in chitin deacetylase genes. Expression profiling of the biotrophic stage of the coffee rust *Hemileia vastatrix* suggested that two chitin deacetylases were most strongly expressed in the early stages of host invasion, coinciding with spore germination and tissue invasion (Vieira et al., 2012). Likewise, the rust fungus *Uromyces viciae-fabae* exhibits massive chitin deacetylase activity when the fungus starts to penetrate through the stomata (Deising and Siegrist, 1995). An overall view of the data suggests complex regulation of GHs during fungal development to achieve a balance between the degradation of plant cell wall polysaccharides without triggering an immune response, and remodeling of the fungal cell wall via both degradation and synthesis to allow growth and development.
Chapter 3: Transcriptome analysis of Pst

Fungal GHs could potentially also play roles in anti-host defence. There is evidence of deposition of host-derived polysaccharides such as arabinogalactan proteins, hydroxyproline-rich glycoproteins, cellulose, callose, pectin, and fucosylated xyloglucans, into the extrahaustorial matrix of old monokaryotic haustoria of the rust fungus *Uromyces vignae* on its host-*Vigna sinensis* (Mendgen et al., 1995). Similarly, encasement of the haustoria of the coffee rust fungus *Hemileia vastatrix* with callose, cellulose, and pectins was detected in the advanced stages of infection of coffee plants (*Coffea arabica*), which however occurred too late to prevent fungal growth and sporulation (Silva et al., 1999). Haustorial encasement structures are associated with host defence in monocot and dicot plants when challenged by non-adapted fungal and oomycete pathogens (Heath et al., 1996, Soylu et al., 2004). A study of a non-adapted isolate of the powdery mildew *Golovinomyces cichoracearum* on Arabidopsis showed that post-invasion non-host resistance is mainly manifested as callosic encasement of the haustorial complex, and the hypersensitive response (Wen et al., 2011). The same phenomenon was described for *Pst* on the non-host broad bean (Cheng et al., 2012). Thus, adapted pathogens must have evolved the ability to suppress or degrade the encasement in order to sustain functional haustoria. Overall, GHs could play fundamental and diverse roles during infection, for example for entry into plant tissue, for fungal wall maintenance, or alternatively for preventing encasement by the host.

### 3.3.5.2. Transporters

Nutrient acquisition is presumed to be a major role of haustoria in rust infection. I therefore searched for potential nutrient transporters by BLASTx of the *Pst* contigs set against the NCBI non-redundant protein and the Transport Classification databases (TCDB: http://www.tcdb.org). The initial cut off for evaluation was e-value <10^-25, and the hits were further annotated manually to remove alignments of less than 100 amino acid residues. The major transporter classes that I found constituted those for amino acids and oligopeptides, sugars, small molecules and ions, and the vitamin co-factor nicotinic acid. Some of these were expressed differentially in a manner that gives insight to the pathogenic strategy of the fungus (Table S7, Transporters).

#### 3.3.5.2.1. Amino acid and oligopeptide transporters

Pathogenic fungi have large requirements for nitrogenous compounds for macromolecule biosynthesis. The best characterized haustorial amino acid transporters are from *U. fabae*: AAT1p
Chapter 3: Transcriptome analysis of Pst

I found very close \textit{Pst} homologs to all three transporters (PST79_2586, PST79_1706, PST79_2986), each of which was expressed at very low levels in germinated spores, but highly in haustoria, especially the AAT2 homolog. I also found three other putative amino acid transporters, one highly similar to AAT3 (PST79_3074) that was expressed strongly in germinated spores but at low levels in haustoria. Similarly to \textit{U. fabae} (Voegele and Mendgen, 2003), my data suggest that most amino acid uptake occurs via haustoria, but the high expression of one amino acid transporter in germinated spores might suggest that uptake can also occur from other cell types such as infectious hyphae. It would be very interesting to test the specificity and affinity of \textit{Pst} putative amino acid transporters; presumably the early expressed transporter should show very high affinity because of the scarcity of nutrients in the apoplast. In addition to this, eight predicted oligopeptide transporters (OPT) were found in my data, three up-regulated in haustoria (PST79_9468, PST79_1656, PST79_8271), one in germinated spores (PST79_2988) and the other four did not show expression biases. The amplification of the OPT gene family in \textit{Pgt} and \textit{Mlp} was proposed to be a genomic adaptation of these pathogens to obtain amino acids, nitrogen and carbon from oligopeptides derived from their host (Duplessis et al., 2011). Overall my data suggest that haustoria are more active than germinated spores in uptake of amino acid compounds, although the early expression of some of those transporters could be crucial for the pathogen’s development.

3.3.5.2.2. ABC and MFS transporters

ABC transporters constitute a large superfamily of primary active transport systems that are present in all kingdoms of life, and play diverse physiological roles in trafficking of a wide range of substrates across internal and external membranes. They represent one of the largest and most ancient transporter classes and derive their name from the shared highly conserved domain, the ATP binding cassette (ABC), which binds and hydrolyzes ATP (Rees et al., 2009). MFS transporters are an ancient class of single-polypeptide secondary carriers which transport small solutes in response to chemiosmotic ion gradients (Pao et al., 1998). I identified twelve transcripts encoding proteins with similarity to ABC transporters in the \textit{Pst} transcriptome, but only three of these showed differential expression (PST79_11360, PST79_10409, PST79_132, up-regulated in germinated spores), so no generalizations about biological function were possible. In other biotrophic pathogens such as \textit{Blumeria graminis} f.sp. \textit{hordei}, a marked reduction in the number of genome-encoded ABC transporters (only 20, equivalent to 50% of the average number in Ascomycete species (Kovalchuk et al., 2013)) has been correlated with the loss of secondary metabolic enzymes, reported as the lowest
number known in fungi (Spanu et al., 2010). In contrast, ABC transporters appear to play important roles during pathogenesis for the hemibiotroph *M. oryzae*, especially in the appressorial stage where far more ABC transporter genes are expressed than in any of the studied stages of *Pst* (Sun et al., 2006, Gupta and Chattoo, 2008, Soanes et al., 2012). This may reflect a need for the pathogen to exclude small molecule defence compounds secreted by the host, and to deploy secondary metabolites during tissue colonization (Soanes et al., 2012). I also identified twenty six transcripts annotated as MFS-like transporters, 16 with putative substrates, and ten classified as general MFS transporters. Of the general MFS transporters, six were expressed preferentially in germinated spores (PST79_7841, PST79_11441, PST79_3438, PST79_8377, PST79_3969, PST79_87060), and two in haustoria (PST79_96, PST79_7181). Duplessis et al (2011) reported 51 and 88 MFS transporter genes in the *Pgt* and *Mlp* genomes respectively, however while the total number of such transporters in the *Pst* genome is presumably similar, only a few were expressed during the developmental stages and sampling conditions studied here. Nevertheless, most of the MFS transporters in *Pst* were preferentially expressed in germinated spores, which is the first infectious stage that encounters host defence. This could suggest that although rust pathogens are also equipped to export toxins or metabolites, the most important roles of these proteins is in growth and development.

### 3.3.5.2.3. Sugar transporters

In order to grow, a biotrophic fungus must secure an organic carbon source from the plant. In contrast to necrotrophs, biotrophic fungi do not secrete a large diversity of extracellular enzymes that depolymerise polysaccharides. Instead, these pathogens set up a specialised interface to allow sequestration of nutrients directly from host cells (Talbot, 2010). The uptake of sugars, which are abundant within plants and a direct product of photosynthesis, is one of the goals of fungal pathogens. During the infection, the pathogens must respond to changes in the availability of these carbon sources by establishing different transport systems in the plasma membrane (Pereira et al., 2013). In my analysis of the *Pst* transcriptome three putative glucose transporters (MFS), seven unspecified sugar transporters (MFS), one mannose transporter and one oligosaccharide translocation protein were identified. One of the transcripts classified as a glucose transporter in *Pst* (PST79_113) shares 92% similarity at the amino acid level with the hexose transporter HXT1p of *U. fabae* (Voegele et al., 2001). This gene showed the highest haustorial expression of all of the sugar transporters identified, and similarly to *Uf*-HXT1, showed almost no expression in germinated spores. The expression pattern of *Uf*-HXT1p is very similar to the amino acid transporter *Uf*-AAT2p.
Chapter 3: Transcriptome analysis of Pst (Voegele et al., 2001). Interestingly, the haustorial expression values for the Pst homologs of these two transporters were also almost identical, suggesting that expression of these genes is coordinated in both rust species. I also identified three Pst homologs of the H⁺-ATPase transporter Uf-PMA1 (PST79_1028, PST79_343, PST79_4481) which is believed to be important for establishment of a proton gradient for coupling substrate translocation into haustoria (Voegele and Mendgen, 2011). Similarly to Uf-PMA1, all of the Pst homologs were expressed more strongly in germinated spores than haustoria (Struck et al., 1998). Biochemical characterization of this U. fabae ATPase showed that despite the higher transcript levels in germinated spores than in haustoria, its enzymatic activity is far greater in haustoria where it may act as a proton-substrate symport mechanism for Uf-HXT1p (Struck et al., 1996, Struck et al., 1998). Thus, this mechanism might be conserved in Pst and potentially outlines a conserved transport mechanism for glucose that is key to the biotrophic lifestyle of rust fungi.

The remaining two putative glucose transporters (PST79_2057, PST79_2930) were up-regulated in germinated spores, suggesting that the fungus prepares for tissue invasion and is poised to exploit host carbohydrates as a nutritional source, or alternatively that these could export substrates for cell wall synthesis. M. oryzae expresses sugar transporters at an analogous stage; during appresoria formation but before tissue penetration, supporting my reasoning (Soanes et al., 2012). Finally, of the seven unspecified sugar transporters, three were up-regulated in haustoria and three were more strongly expressed in germinated spores. One of the genes expressed in germinated spores (PST79_2314) has a high similarity to STL1 (e-val 1x10⁻¹⁰¹ with the Saccharomyces cerevisiae gene), a glycerol-proton symporter that is inactivated by glucose (Ferreira et al., 2005). This might suggest that glycerol could be a primary carbon source, or alternatively plays a role as an osmoprotectant during spore germination, as has been shown for other fungi (Ansell et al., 1997, Castro and Loureiro-Dias, 1991, Wei et al., 2004, Scarpari et al., 2005).

3.3.5.2.4. Phosphate transporters

Inorganic phosphate is an essential nutrient required for the synthesis of nucleic acids, phospholipids, cellular metabolites, and protein modification, reactions that require high phosphate concentrations (Wykoff and O'Shea, 2001). I identified four transcripts encoding different phosphate transporters from my dataset. One of these was annotated as the phosphate translocator of the inner mitochondria membrane, discussed below. From the remaining three, two were up-regulated in germinated spores (PST79_3467, PST79_11745) and annotated as inorganic phosphate
transporters related to Pho84, a phosphate starvation-inducible high-affinity $\text{H}^+/\text{PO}_4^{3-}$ symporter (Bun-Ya et al., 1991). The third transporter (PST79_8877) was annotated as an SPX-domain containing protein, a domain often present in proteins involved in regulating inorganic phosphate homeostasis (Secco et al., 2012). In budding *Saccharomyces cerevisiae* cells, response to phosphate starvation includes the up-regulation of the activity of a high-affinity phosphate uptake system (Nieuwenhuis and Borst-Pauwels, 1984). Pho84p is one of the key proteins in this phosphate transport system, which is regulated transcriptionally in response to extracellular phosphate levels sensed by a phosphate-responsive signal transduction pathway (Lenburg and O'Shea, 1996). The expression of this gene is greatly induced in low-phosphate medium and is essential for growth under low-phosphate conditions (Bun-Ya et al., 1991). The presence of a homologous gene in *Pst* and its up-regulation in germinated spores could suggest that similar phosphate starvation sensing/signalling mechanisms exist in rust fungi, which could be crucial for cell cycle progression and growth as has been shown in yeast (Popova et al., 2010). In addition, thirteen transcripts similar to putative ion transporters with specificities for $\text{Ca}^{2+}$, $\text{Cu}^{2+}$, $\text{Fe}^{2+}$, $\text{Mg}^{2+}$, $\text{Mn}^{2+}$, $\text{K}^+$ and $\text{Zn}^{2+}$ (PST79_3206, PST79_9341, PST79_4284, PST79_6929, PST79_9148, PST79_10399, PST79_4745, PST79_6508, PST79_11667, PST79_1137, PST79_8091, PST79_5696, PST79_4204) were identified in my *Pst* data. Transporters for $\text{Zn}^{2+}$ and $\text{Mn}^{2+}$ were up-regulated in haustoria, while transporters for $\text{K}^+$ and $\text{Ca}^{2+}$ were up-regulated in germinated spores.

### 3.3.5.2.5. Nicotinic acid transporters

Nicotinic acid is an essential cofactor for many enzymic processes. I found four transcripts (PST79_362, PST79_7845, PST79_4348, PST79_3107) encoding proteins with similarities to the *S. cerevisiae* nicotinic acid transporter TNA1 (e-value $\sim1\times10^{-30}$) (Llorente and Dujon, 2000) in my data, one of them notably up-regulated in haustoria. Vitamin B3 (nicotinamide and nicotinic acid) is essential to living cells because it can be converted to nicotinamide adenine dinucleotide (NAD$^+$), which can be reduced to NADH or the phosphorylated form NADPH. Generally, NADPH is a reductive intermediate for biosynthesis pathways, whereas NADH is an energy precursor and protects against oxidative damage. Thus uptake of host-derived niacin could be important in the haustorial stage which is greatly involved in energy production, biosynthesis of molecules for spore formation, and protection against oxidative stresses. Duplessis et al. (Duplessis et al., 2011) reported the *in planta* expression of four *Mlp* genes with similarity to *S. cerevisiae* TNA1 permease at 96 hpi, supporting the idea that uptake of host nicotinic acid plays an important role during the advanced stages of the infection. I also searched for contigs that encode enzymes required for *de novo* synthesis of niacin
from tryptophan. The known pathway in organisms that use tryptophan as a source for \( \text{NAD}^+ \) (including yeasts, humans and some bacteria) involves six enzymes; I was able to identify five of these in my dataset, two of which were up-regulated in haustoria (Table S3, General metabolism). Thus it remains an open question whether \( \text{Pst} \) can synthesize niacin \textit{de novo}, or it relies exclusively on its host for obtaining this essential cofactor.

### 3.3.5.3. Primary metabolism

#### 3.3.5.3.1. Energy use - oxidative phosphorylation

Transcripts classified in this category were represented more in haustoria than in germinated spores (Figures 3.6 and 3.7). Transcripts with similarity to genes involved in primary pathways of energy production comprising glycolysis, citric acid cycle, and oxidative phosphorylation were mostly up-regulated in haustoria (Table S5, Energy metabolism). Similar observations have been made for the obligate biotrophs \( \text{U. fabae} \) (Jakupovic et al., 2006) and \( \text{B. graminis} \) (Both et al., 2005a) where genes related to glycolysis are up-regulated during the parasitic stage. Those few \( \text{Pst} \) genes with higher expression in germinated spores corresponded to the key enzymes of the glyoxalate cycle, which will be described later. In addition, a phosphate translocator of the inner mitochondria membrane (PST79_4938), which carries out the coupled transport of \( \text{H}_2\text{PO}_4^- \) and \( \text{H}^+ \) for subsequent ATP synthesis via the electron transport chain, was massively up-regulated in haustoria, consistent with a high demand for ATP in this tissue. This could be a consequence of the different nutritional status of haustoria and germinated spores. While the stored energy resources are limited in spores, haustoria can derive sugars abundantly from the host to feed energy production pathways and drive biosynthetic pathways for subsequent development (Figure 3.9).

#### 3.3.5.3.2. Fatty acid metabolism

Acetyl-CoA is the basic currency of carbon metabolism within the cell. It is the immediate product of carbohydrate catabolism and \( \beta \)-oxidation of fatty acids, and when in excess, can be used for biosynthetic processes through the glyoxylate cycle, gluconeogenesis and glyceroneogenesis pathways. In my transcriptomic study, the gene for triacylglycerol lipase (PST79_8009) together with nearly all those for the enzymes of glyoxylate cycle, gluconeogenesis and glyceroneogenesis were significantly up-regulated in germinated spores compared with haustoria (Table S3, General metabolism). This suggests that the activity of metabolic pathways for converting fatty acids into
carbohydrates is important in the initial stages of infection. Key genes involved in β-oxidation of fatty acids such as acyl-CoA dehydrogenase (PST79_1352) and enoyl-CoA hydratase (PST79_4902) were also up-regulated in germinated spores, suggesting that once triacylglycerides are broken down, the released fatty acids are oxidised and the acetyl-CoA produced can be fed into gluconeogenesis by means of the glyoxylate cycle. The two key enzymes in the glyoxylate cycle, isocitrate lyase (PST79_6444) and malate synthase (PST79_4401), were highly up-regulated in germinated spores. However, malate dehydrogenase and citrate synthase, two other enzymes in this cycle, did not show the same biased expression pattern between the two pathogenic stages. Malate dehydrogenase participates in the glyoxylate cycle and is also involved in the malate-aspartate shuttle, which is responsible for translocating electrons produced in glycolysis through the mitochondrial membrane for oxidative phosphorylation. Thus, the role of this enzyme is expected to be important in germinated spores but also in haustoria because of their high glycolytic activity. Apart from fatty acids, the second product of triacylglycerol breakdown is glycerol. This compound can be converted to intermediates of gluconeogenesis by means of glycerol kinase (PST79_6140) and glycerol-3-phosphate dehydrogenase (PST79_3414), genes which also showed significant upregulation in germinated spores. As mentioned previously, a gene similar to STL1 (encoding a glycerol transporter) was also up-regulated in germinated spores, which suggests that glycerol derived from the host apoplast could be taken up and incorporated via the gluconeogenic pathway. Finally, genes encoding enzymes involved in glycogen breakdown were also up-regulated in germinated spores, which also results in glycerol production. In Pgt uredospores, as well as spores of other plant fungal pathogens, glycogen is a stored energy source that is utilized during germ tube extension (Ehrlich and Ehrlich, 1969, Thines et al., 2000). These observations are consistent with data from other biotrophic and hemibiotrophic fungi such as B. graminis and M. oryzae (Both et al., 2005a, Soanes et al., 2012), where similar metabolic activity has been detected in early stages of the infection. Overall, my results point towards spore germination utilising stored compounds including lipids and glycogen as sources for energy production (Figure 3.8), whereas the massive upregulation of HXT1 and enzymes of glycolysis, TCA cycle and oxidative phosphorylation in haustoria suggest that host-derived glucose is the primary carbon source.
Figure 3.8. Metabolic processes in germinated spores. Metabolic processes or specific enzymatic activities overrepresented in spores identified by B2G analysis and manual annotation are highlighted in this cartoon. Orange boxes and yellow ovals are metabolic processes or particular enzymes that showed statistically significant upregulation in germinated spores.

### 3.3.5.3.3. Pentose phosphate shunt (PPS)

The PPS is an alternative fate for glucose which produces NADPH and ribose-5-phosphate instead of entering glycolysis. Ribose-5-phosphate can be used for nucleotide synthesis (oxidative branch), or recycled for energy production via glycolysis (non-oxidative branch). An important product of the PPS is the reductant NADPH, which is used almost exclusively in anabolic pathways, for example biosynthesis of fatty acids, amino acids and nucleotides including ATP. I found evidence for expression of all enzymes of the PPS in both germinated spores and haustoria (Table S4, General metabolism). Interestingly, all of the enzymes participating in the oxidative branch of this pathway showed an expression bias in haustoria as did transaldolase, which belongs to the non-oxidative branch. In germinated spores, the genes encoding the enzymes ribulose-phosphate 3-epimerase (PST79_10575) and phosphoglucomutase (PST79_5659) were up-regulated. The ability of this
pathway to work in different directions suggests that it could play different roles in each pathogenic stage. While both pathogenic stages are expected to have a strong requirement for biosynthetic processes, this is expected to be higher in haustoria which presumably contribute to the provision and transport of substrates to fuel the growing infection and eventual spore biogenesis (Figure 3.9). Thus, the haustorial PPS appears to engage the oxidative arm for maximal NADP$^+$ reduction and ribose-5-phosphate production. Additionally, an important enzyme of the non-oxidative branch, transaldolase, showed higher expression in haustoria than in germinated spores. The non-oxidative branch can convert five carbon sugars into six and three carbon sugars, which are fed into glycolysis. This would appear to be unnecessary in haustoria which should not be limited for glucose. Alternatively, the non-oxidative arm of PPS may work in the opposite direction, taking fructose-6-phosphate and glyceraldehyde-3-phosphate generated by glycolysis to produce ribose-5-phosphate, which is the basis of production of ATP, NAD$^+$ and nucleic acids. However, the remainder of the enzymes that participate in the non-oxidative arm did not show a clear expression trend, which precludes clear conclusions on the role of this branch of the PPS. In Plasmodium falciparum, a human pathogen with an intracellular niche that is comparable to the rust haustorial stage, the oxidative arm is thought to operate during early stages of parasite development while the non-oxidative arm is active later in the cycle (Bozdech and Ginsburg, 2005, Bozdech et al., 2003). These studies also showed that transcription of the PPS genes is not always coordinated, with activity of the pathway as a whole often dependent on expression of a single gene whose transcription is the rate limiting step for deployment of the pathway’s activity. In B. graminis, the absence of clear coordination in expression of the enzymes of PPS has also been observed (Both et al., 2005a). These discrepancies require direct biochemical investigation which will provide more clues about the physiological implications of these patterns.
Figure 3.9. Metabolic processes in haustoria. Metabolic processes or specific enzymatic activities overrepresented in haustoria identified by B2G analysis and manual annotation are highlighted in this cartoon. Orange boxes and yellow ovals are metabolic processes or particular enzymes that showed statistically significant upregulation in haustoria. Light orange boxes are metabolic processes where some genes showed statistically significant upregulation in haustoria, and others showed a tendency to be more expressed in haustoria but were not statistically significant.

3.3.5.3.4. Thiamine biosynthesis

A number of metabolic enzymes require the vitamin B1 derivative thiamine pyrophosphate as a cofactor. Examples from the pathways described here include alpha–ketoglutarate dehydrogenase (α–KGDH) in the TCA cycle; transketolase in the non-oxidative PPS; and pyruvate dehydrogenase (PDH) which connects glycolysis and the TCA cycle. I found that transcripts associated with the thiamine biosynthesis pathway were massively up-regulated in haustoria, with the exception of the genes for thiamine-phosphate kinase and thiamine pyrophosphokinase which could not be found in my data. These enzymes are responsible for phosphorylating thiamine phosphate and thiamine diphosphate respectively, and interestingly, the latter is also absent in the genome of the malarial parasite *P. falciparum* (Bozdech and Ginsburg, 2005). The genes I found include those encoding homologs of thiamine biosynthesis genes *TH1* (4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate synthase)(PST79_1423, PST79_612, PST79_1422) and *TH2* (hydroxyethylthiazole phosphate synthase)(PST79_882) from *U. fabae* identified previously by Hahn and Mendgen in 1992.
The massive expression of these genes almost exclusively in haustoria agrees with studies done in other rust fungi (Duplessis et al., 2011, Thara et al., 2003), but it is not clear why the pathogen requires such high expression at this stage, which appears excessive if the role is simply to produce sufficient cofactors for metabolic enzymes. One explanation may be that thiamine is exported to other tissues. Alternatively, thiamine is known to alleviate stress in different organisms as an antioxidant. It is interesting that in the oomycetes, all haustorium-forming species have lost the thiamine biosynthetic pathway (Kemen et al., 2011), which suggests that they must obtain thiamine from the host.

3.3.5.3.5. Nitrogen metabolism

Previous genome studies of obligately biotrophic fungal plant pathogens have noted the absence of genes encoding enzymes for nitrate and sulfur assimilation (Duplessis et al., 2011, Spanu et al., 2010). Likewise, I was unable to identify genes for nitrate reductase or nitrite reductase, or for a nitrate transporter that could import nitrate directly from host tissue. Instead, I found a transcript encoding a protein with very high similarity to an ammonium transporter (PST79_5204) expressed in both germinated spores and haustoria. Ammonium is the preferred form of nitrogen for most organisms (Meti et al., 2011) and is assimilated into glutamate and glutamine. Glutamate synthase and glutamine synthetase are the key enzymes involved in ammonia assimilation. Glutamate is formed from α-ketoglutarate and glutamine in a reaction that is catalyzed by glutamate synthetase, followed by amination of glutamate by NH4+ to form glutamine catalyzed by glutamine synthetase. Subsequently, glutamate is used to synthesize the majority of amino acids, whereas glutamine serves as an amino-donor during other biosynthetic processes. Interestingly, the transcripts encoding genes for both glutamine (PST79_3032) and glutamate synthase (PST79_2575) identified here were up-regulated in haustoria, consistent with a greater role for biosynthesis in this organ. Genes encoding other enzymes that participate in subsequent steps including glutamate dehydrogenase (PST79_536), aspartate aminotransferase (PST79_7929), asparagine synthase (PST79_1116), and asparaginase (PST79_83), were all expressed in both pathogenic stages, suggesting nitrogen compounds obtained from the host can be converted freely into compounds required for pathogen growth and development.

3.3.5.3.6. Sulfur metabolism
Sulfur is an essential component of living cells as it is a fundamental component of the amino acids methionine and cysteine, Coenzyme A, and iron-sulfur enzymes. Most fungi take up sulfur as sulfate which is reduced to sulfide as a precursor of cysteine (Marzluf, 1997). Despite this, the steps for sulfate assimilation are energetically expensive and fungi prefer to utilise cysteine or methione. To date, enzymes for sulfate uptake and reduction have not been identified in obligate biotrophs such as *Hyaloperonospora arabidopsidis* (Baxter et al., 2010), *B. graminis* (Spanu et al., 2010), or *P. graminis* f.sp. *tritici* (Duplessis et al., 2011). In addition, genes encoding enzymes for sulfate reduction including sulfite reductase and phosphoadenosine phosphosulfate reductase were absent, consistent with observations for other obligate plant pathogens. However, I found a contig (PST79_3132) that encodes a transporter with very high similarity to S-methylmethionine (SMM) permease, which was almost exclusively expressed in haustoria compared to germinated spores. This is interesting because SMM is present at very high levels in wheat phloem (Bourgis et al., 1999) and as such, could be an abundant source of sulfur metabolites for the parasite. Future experimentation on this gene will be important to determine the real capability of the corresponding protein to transport SMM, which could constitute a new mechanism for sulfur assimilation. Apart from this, I was unable to identify candidate cysteine or methionine transporters in my data. Lastly, I identified a gene encoding cysteine synthase (PST79_10177), which assimilates reduced sulfur (\(S^2\)) into cysteine, expressed at low levels in both pathogenic stages. Genes with high similarities with those involved in the interconversion of homocysteine and cysteine through the intermediary formation of cystathionine (transsulfuration pathways) were also identified. This pathway creates cysteine from methionine.

### 3.3.5.4. Protection against host defences

#### 3.3.5.4.1. Glutathione metabolism

Glutathione provides redox buffering to cells, and plays a role in many cellular processes including iron metabolism (Pocsi et al., 2004). The anti-oxidant properties of glutathione contribute to anti-defence against host active oxygen compounds. Glutathione is derived from the amino acids glycine, glutamate and cysteine. The two key enzymes for synthesis of glutathione are glutamate-cysteine ligase (PST79_3565) and glutathione synthetase (PST79_1178). Both of these genes showed greater expression in haustoria, and PST79_3565 was significantly upregulated. This could suggest an ongoing need for antioxidants in the static haustorial stage which is exposed to plant defences. I also found that glutathione reductase (PST79_10658) and glutathione peroxidase (PST79_5949)(enzymes
involved in the redox balance), as well as genes involved in the response to oxidative stress, were more highly expressed in germinated spores than in haustoria (Table S4, General metabolism). All of this would suggest the presence of glutathione in spores prior to development. Thus, one function of haustoria could be to synthesize glutathione for subsequent deposition into new spores as a pre-formed environmental defence.

3.3.5.4.2. Cytochrome P450s

Cytochrome P450s comprise a diverse superfamily of proteins containing a heme cofactor that catalyse the oxidation of organic substances. Among their most common substrates are steroidal molecules as well as drugs and toxins. The P450 system is important in fungal evolution for adaptation to ecological niches (Deng et al., 2007), and its potential role in compound detoxification including metabolism of plant defence molecules (Vanetten et al., 1995, Hahn and Mendgen, 1997). A total of 29 and 17 P450s are encoded by the genomes of *M. larici-populina* and *P. graminis* f. sp. *tritici* (Duplessis et al., 2011) respectively. In the current dataset, I identified only eight expressed P450 genes, five of which were up-regulated in haustoria (Table S4, General metabolism). One of the P450s (PST79_8680) I identified is highly similar to CYP51A1, a lanosterol 14 alpha-demethylase, required for conversion of lanosterol to cholesterol. This enzyme is targeted by the commonly usedazole class of antifungal drugs, which occupy the active site of the enzyme inhibiting the production of ergosterol, a component of the fungal cell membrane (Vanden Bossche et al., 1990). Two other P450s (PST79_1792, PST79_1793) up-regulated in haustoria belong to the CYP67 family, which was first reported in *U. fabae* as an in planta-induced gene (Hahn and Mendgen, 1997). The remaining genes encode unspecified mono-oxygenases.

3.3.5.5. Transcriptions factors

Overall, my analysis revealed approximately 37 transcripts encoding proteins related to transcription factor activity, from both haustoria and germinated spores (Figure 3.7 and Table S6, HeatMap). Despite this, it is difficult to obtain specific insights into the patterns and functions of these genes. However, there were a few exceptions where I could predict a possible function based on BLAST2GO and BLASTx searches (Table S4, General metabolism). One of the most interesting genes found in my dataset was a Ste12-like gene (PST79_9215), which was up-regulated in haustoria. Ste12 was first identified in a yeast sterile mutant, and multiple studies have shown that Ste12-like genes play major roles in regulating morphogenetic programs in response to environmental changes (see review
Chapter 3: Transcriptome analysis of Pst

(Yong Sak Hoi and Dumas, 2010). Yeast Ste12 is an important regulator of invasive growth and pseudohyphal development (Madhani and Fink, 1997) and its homologs in a number of fungi are important for sexual development and pathogenicity (Liu et al., 1994, Calcagno et al., 2003, Tsuji et al., 2003, Chang et al., 2001). In the root pathogen Fusarium oxysporum, a ste12 mutant was impaired in pathogenesis on tomato and the pathogen could not differentiate into specialized infectious structures (Rispail and Di Pietro, 2009). In the hemibiotrophic fungal pathogens Colletotrichum lagenarium, Colletotrichum lindemuthianum and M. oryzae, Ste12-like genes are important for penetration of leaf surfaces from appresoria and subsequent invasive growth (Park et al., 2002, Tsuji et al., 2003, Wong Sak Hoi et al., 2007). Moreover, some evidence suggests that Ste12 plays a role in host penetration of the arbuscular mycorrhiza Glomus intraradices (Tollot et al., 2009). Although the role of Ste12 transcription factor in rust fungi is unknown, I suggest that the Pst homolog could play a role in specifying haustorial development.

A second interesting gene was the homolog of TUP1 (PST79_2626), a conserved transcriptional regulator from fungi and mammals, which controls dimorphism in certain fungi (Nadal et al., 2008, Grbavec et al., 1999). The change in morphology between yeast-like growth and a filamentous state in response to environmental signals is frequently associated with virulence in plant pathogenic fungi (Nadal et al., 2008). TUP1 plays a central role in controlling the expression of the genes implicated in the genetic control of mating, filamentation, and pathogenic development of Ustilago maydis (Elias-Villalobos et al., 2011). Analyzing the function of TUP1 in rust fungi could provide a better understand of how it acts within the unique biological context in which these pathogens develop.

Lastly, I found five other genes associated with transcription factor activity in my data, encoding a putative CCAAT-box binding factor (PST79_5498), a putative activator of basal transcription (PST79_5377), an MCM-domain-containing protein (minichromosome maintenance ATPase) (PST79_3174), a putative bZIP transcription factor (PST79_8536) and a gene with similarity to the regulator Cys3 (PST79_492) reported originally in Neurospora crassa (Paietta et al., 1987). Cys3 controls the synthesis of a set of catabolic enzymes for utilization of secondary sulfur sources when sulfur-containing amino acids and inorganic sulfate are missing or limited (Tao and Marzluf, 1998). The presence of a similar gene in Pst might suggest that secondary sulfur sources play a role in biotrophy, and upregulation of this gene in haustoria could suggest these sources are derived from the host.

3.3.5.6. Signal transduction
Cellular differentiation and filamentous growth are finely regulated in many fungi in response to environmental and nutritional signals. External stimuli must be transmitted within the fungus by signal transduction pathways to enable the appropriate response. My transcriptomic analysis revealed high expression of signal transduction components in both germinated spores and haustoria. Transcripts encoding proteins related to two component response regulators, the Ras protein superfamily of guanine nucleotide exchange factors, and diverse protein kinases were most highly represented. In fungi, two-component signalling systems participate in processes such as environmental sensing, oxidative stress response, cell-cycle control, and switching between non-pathogenic and pathogenic states (Bahn, 2008, Oide et al., 2010, Catlett et al., 2003). They are typically composed of a sensor-kinase protein, a phosphorelay protein, and a response regulator. I found genes encoding two-component regulatory proteins in both haustoria and germinated spores (Table S6, HeatMap). Generally, the proteins expressed in each pathogenic stage were distinct, consistent with the idea that they respond to different stimuli. There was a larger representation of these genes in germinated spores suggesting a more complex interpretation of environmental signals prior to development of the haustorial structure. One gene up-regulated in germinated spores (PST79_186) has high similarity to Os-1, a two-component histidine kinase originally isolated from Neurospora crassa. Os-1, also known as Nik-1, is required for adaptation to high osmolarity (Alex et al., 1996, Schumacher et al., 1997). Three groups of selective fungicides target the osmotic stress signal transduction pathway (Cui et al., 2002, Pillonel and Meyer, 1997). Intracellular osmotic pressure plays important roles in fungal development, for example as inside the M. oryzae appresorium (Howard and Valent, 1996). Although disruption of HIK1 (Os-1) in this pathogen caused no defect in growth on normal media or in pathogenicity to rice plants, the mutant strain acquired resistance to three groups of fungicides (phenylpyrroles, dicarboximides, and aromatic hydrocarbons) (Motoyama et al., 2005). The almost exclusive expression of this gene in germinated Pst spores suggests that the osmosensitive signal transduction pathway could play a role in the early stages of stripe rust disease.

Ras proteins are a superfamily of guanine nucleotide exchange factors in which the phosphorylation status of the bound nucleotide operates a binary switch. They have numerous signaling roles in cellular functions including cytoskeletal integrity, proliferation, differentiation, cell adhesion, and cell migration. Ras proteins are active in the GTP-bound form, which is promoted by guanine nucleotide exchange factors (GEF) in an exchange reaction with GDP, and deactivated by GTPase activating proteins (GAP) that promote dephosphorylation of the bound GTP nucleotide. I identified evidence of expression of genes encoding proteins with similarity to Ras proteins called ADP ribosylation
Chapter 3: Transcriptome analysis of Pst

3.4. CONCLUSION

In this study, a combined long-read/short-read sequencing strategy resulted in an effective way to collect gene expression data when an annotated genome sequence is not available. While the 454 platform gives significantly longer reads which helps with transcript assembly, shorter Illumina reads give the density of coverage necessary for robust statistical comparisons. One important aspect of defining transcriptomes, as described here, is that it reduces the complexity of the genome. Thus, while the predicted coding capacity of rust genomes is ~22,000 genes, I identified ~12,000 expressed contigs in the sampled tissues. Presumably, many of the remaining genes are expressed in different contexts (for example on the alternate host, or as part of different spore stages), or were simply undetected or unassembled in my study.

Another important aspect of my study is quantification of expression using Illumina data, which considerably enhances the accuracy and reliability of my expression measurements. Thus, use of 454 and Illumina together provides a robust expression dataset. The current data provide a comprehensive view of the stripe rust transcriptome in germinated spores and haustorial stages. Overall, the fungus appears to use similar biochemical pathways to those described in better characterised species such as *M. oryzae*, but with adaptation to the particular pathogenic lifecycle of wheat rust pathogens. Haustoria and spores are clearly working in different ways. The germinating spores prioritises polarised growth from stored energy until such time as it is able to mine its own resources, whereas haustoria are a sessile stage devoted to nutrient extraction and defence suppression. The uredospore contains vast stores of energy in the form of lipid bodies, sugar alcohols and proteins (Staples and Wynn, 1965, Harder, 1984) to allow the foraging germ tube to find the pathogenic niche, and in that sense is analogous to a plant seed. From the data, it appears that haustoria can create abundant demand for sugars, amino acids, and nitrogen, through high expression of transporters for these fundamental building blocks. In addition, the uptake of methylmethionine appears to be an alternative mechanism for sulphur assimilation. Sugars are used immediately by haustoria for the production of ATP through glycolysis, TCA cycle and oxidative phosphorylation. By contrast, spores appear to obtain energy by utilising lipid resources via the
glyoxylate pathway. A further important haustorial function is biosynthesis of macromolecules, and it is clear that steady production of the reductant NADPH is made possible by import of nicotinic acid through dedicated transporters, and its ultimate conversion to NADPH by the PPS. Spores also use the PPS but it is likely that this is works in the opposite (oxidative) direction, for energy production. Molecules made in haustoria are presumably exported directly to spores for biogenesis and storage, so the molecular pathways active in haustoria should reflect spore composition directly. With respect to the completeness of my dataset, I note as have previous authors that many genes are novel so cannot be annotated. For this reason, it is difficult to exclude the presence of certain genes with major implications for the biotrophic lifecycle, such as N and S assimilation. Of course, the absence of these genes could also be ascribed to low expression of mis-assemble of small reads. These are important challenges for the field because barring technological breakthroughs, most fungal genome sequences are likely to have large components of small sequence reads in their genesis.
Chapter 4

Purification of *Pst* haustoria

4.1. INTRODUCTION

Rust fungi are completely dependent on a living host to grow and reproduce. In the case of wheat stripe rust, the infection starts when the uredospore germinates and enters through the cuticle or through stomata of the leaf, eventually forming an infection structure called the appressorium (Kolmer et al., 2009). Much of the ultrastructural work on the cereal rusts has provided a comprehensive view of the physical events of the infection process (Bushnell, 1972, Bushnell and Roelfs, 1984). Once inside the leaf, the formation of other complex infection structures including, as the substomatal vesicle, infectious hyphae, haustorial mother cells, haustoria and sporogenic structures allows the pathogen to complete the invasion of its host. From the haustorial mother cell, the rust fungus penetrates the host cell wall and differentiates the haustorium in such a way that although it is completely enclosed by the plant cell wall, the cytoplasm of the host and parasite cells remain separate, Figure 4.1 (Voegele et al., 2009, Bushnell, 1972). The membrane separating the haustorium from the host cytoplasm is called the extrahaustorial membrane (EHM), which in turn is
Chapter 4: Purification of Pst haustoria

110

separated from the haustorium by the extra haustorial matrix (EHMx) (Harder and Chong, 1984). EHMx is thought to contain material from fungal and plant origin (Harder and Chong, 1991), and cytochemical evidence indicates that this matrix is mainly composed of carbohydrates and glycoproteins (Hippe-Sanwald et al., 1994, Mendgen and Hahn, 2002, Stark-Urnau and Mendgen, 1995). The haustorium is connected to the fungal mycelium by a narrow junction called the haustorial neck which in several ultrastructural studies have been shown is surrounded by a neck band (Harder and Chong, 1991). This band is thought to serve a role in sealing of the EHM from the bulk apoplast and the plant plasma membrane (Panstruga, 2003). The development of the haustorium is probably essential for successful infection as this structure appear to play an active role in the transport of nutrients from the host, and their subsequent metabolism into energy for fungal growth (Garnica et al., 2013). It is also believed to have an active role on the secretion of effector proteins (Garnica et al., 2013), which can impact pathogenicity in both positive and negative ways (Voegele and Mendgen, 2003).

**Figure 4.1. Generalized illustration of infection structures of rust fungi.** The figure illustrates where the haustorium sits in relation to the host cell. The haustorium develops from a haustorial mother cell and is formed between the cell wall and a plasma membrane, separating it from the host cytoplasm. Fungal structures: HMC, haustorial mother cell; EHMx, Extrahaustorial matrix; PM, Fungal plasma membrane.

Dissection of the activities of haustoria has been hampered by the difficulty in obtaining these cells at purity in significant numbers. Usually chloroplasts are the main contaminant, far outnumbering haustoria cells. Secondly, as an obligate biotroph, only partial differentiation of infection structures can be induced *in vitro* (Wietholter et al., 2003, Dickinson, 1969, Heath and Perumalla, 1988, Hoch
Chapter 4: Purification of Pst haustoria

and Staples, 1987, Read et al., 1992). The induction of haustorial differentiation in vitro has been reported, although these cultures tend to show aberrant genotypes and possibly of uncertain gene content (Heath, 1989), being unsuitable for studying biotrophy. The first significant step in overcoming this problem was made by Hahn et al (1992) who developed a chromatography technique to isolate haustoria from Uromyces viciae-faba based on affinity of the lectin Concanavalin A for isolated haustoria (ConA affinity method, see Chapter 2). The technique works for other rust pathogens including M. lini, Pgt and Pst, however its efficiency varies enormously from one pathogen to another (Garnica et al., 2013, Hahn and Mendgen, 1992, Catanzariti et al., 2006). In contrast to rust fungi, haustoria from the powdery mildews Sphaerotheca fuliginea, Erysiphe pisi and B. graminis could not be isolated by this method, presumably because their EHMX differs in composition from those of of rust fungi (Hahn and Mendgen, 1992). The contamination of haustorial preparations with chloroplasts can be a significant disadvantage of this method if haustorial purity is necessary for subsequent experimentation. In previous Chapters it was shown that the application of this technique to Pst allows the isolation of approximately 6.0 x 10^5 haustoria per 15 g of infected tissue with a haustoria:chloroplast ratio of approximately 1:15. This is about 62-fold less haustoria than when the same technique is applied to Uromyces viciae-fabae (Hahn and Mendgen, 1992).

A variety of methods for purification of inter- and intracellular infection structures of rust, powdery mildew and downy mildew fungi from infected plant tissues by density gradients have been devised. These include gradients based on glycerol (Gil and Gay, 1977, Mackie et al., 1991), sucrose (Crucefix et al., 1987, Tiburzy et al., 1992, Song et al., 2011), Ficoll (Cantrill and Deverall, 1993), and Percoll (Pain et al., 1994, Micali et al., 2011). In these methods, the infected tissue is typically homogenized in a suitable buffer, before layering of the homogenate onto a gradient of one of these media. During centrifugation, the various particles move through the gradient and are separated solely on the basis of their density. The particles appear as bands or zones in the gradient, facilitating their concentration and subsequent recovery. Although the efficiency of these methods and the final degree of contamination varies from one protocol to another, density gradients are flexible and can be adjusted according to the biological sample. Percoll, an inert colloidal silica, appears to have become the medium of choice since it fulfills almost all the criteria of an ideal density gradient medium (Pertoft, 2000). One important characteristic is that it allows the formation of isosmotic gradients, preventing cell death by osmotic stress. The versatility of Percoll offers a wide range of methods for density purification or enrichment of cell populations before attempts at higher resolution separation methods. To date, density gradients have not been used to isolate Pst haustoria.
In an important technical advance, Takara et al. (2009) showed that a combination of Percoll gradients and flow cytometry was effective to pre-enrich and then purify primary hyphae from the plant pathogen *Colletotrichum higginsianum*. The specific selection of viable hyphae using the vital dye fluorescein diacetate effectively eliminated contamination by other fungal cell types and plant components. The hyphae were recovered in a viable condition with a purity of 94%, allowing the authors to construct a stage-specific cDNA library, which led to the identification of three fungal genes which are likely to encode effector proteins. This example, together with others on mammalian cells, (Wong et al., 1991, Goncalves et al., 2007, Campanella et al., 2002, Lassalle et al., 1999), points to Percoll-flow cytometry as an attractive alternative to purify fungal haustoria. Flow cytometry is a laser-based technology that measures optical and fluorescence characteristics of single cells or any other particle, including subcellular organelles, nuclei, and chromosome preparations (Brown and Wittwer, 2000). Fluorescence-activated cell sorting (FACS) is a specialized type of flow cytometry which allows the sorting of a heterogeneous mixture of biological cells into two or more populations based upon the specific light scattering and fluorescence characteristics of each cell type. The fluorescence of the cells or subcellular particles of interest can manipulated by tagging them with a specific antibody(ies) linked to a fluorescent dye, or other methods of introducing internal or external fluorescent tags, or simply by using autofluorescence properties. The technology is useful for cell counting, cell sorting, and biomarker detection amongst other applications. Physical properties, such as size and internal complexity or fluorescent properties can resolve certain cell populations (Brown and Wittwer, 2000). Simultaneous multiparametric analyses are possible, making this technology a valuable technique for analysis of individual cells within complex populations (McCoy, 2002, Mattanovich and Borth, 2006). Plant tissues possess high levels of autofluorescence from chlorophyll in chloroplasts (Zhang et al., 2010). Chlorophyll absorbs in the blue region of the visible spectrum, which produces a significant amount of fluorescence at wavelengths >600 nm when excited with wavelengths between 420 and 490 nm (Rowan et al., 2007, Fricker and White, 1992). This characteristic makes fluorescent particles of chloroplasts and can be used as a tag to isolate and separate them from other particles, for example haustoria labeled with fluorescent lectins.

In this chapter, I describe how the physical qualities of *Pst* haustoria were exploited to develop an efficient two-phase method for their enrichment and purification. First, discontinuous and continuous Percoll density gradients were investigated as alternative techniques to the ConA-affinity method to generate enriched suspensions of stripe rust haustoria. Second, the application of FACS
to haustoria labeled with a fluorescent lectin was investigated as a way to effectively eliminate contamination by plant components. The processes to develop both phases and the results of combining them are described in detail.

4.2. MATERIALS AND METHODS

4.2.1. Haustoria enrichment on Percoll density gradients

4.2.1.1. Processing of *Pst*-infected tissue prior to density gradient centrifugation

The collection and processing of infected tissue was undertaken as described in Chapter 2, Sections 2.2.1 and 2.2.2.2. Between 15-20 g of infected wheat leaves (9 dai) were collected, washed and sterilized with bleach, before further washing with water and 70% ethanol. Ethanol was removed by washing with water and the infected tissue was dried. All subsequent steps were carried out at 4°C or on ice. Infected plant material was homogenized using a Waring blender in homogenization buffer [0.2 M sucrose, 20 mM MOPS pH 7.2] (1x IB), and passed consecutively through 100 µm and 20 µm meshes. The filtrate was centrifuged and the resulting pellets were resuspended in 20 ml of cold 1x IB. This suspension is referred to as the homogenate of infected tissue.

4.2.1.2. General steps for the preparation of discontinuous Percoll gradients

To perform discontinuous Percoll gradients, Percoll (GE Healthcare) solutions of different concentrations were prepared from 100% Percoll, 10x IB and sterile MilliQ water in proportions adjusted to obtain the desired final Percoll concentration (% v/v) and an IB concentration of 1x. One ml wide-bore tips and a 1 ml micropipette were used to dispense the Percoll solutions into Nalgene thin wall polypropylene tubes (round-bottom) of 40 ml capacity. The discontinuous gradients were prepared accordingly to requirements by sequentially layering the Percoll solutions from the most concentrated to the most dilute, avoiding mixing between the layers. Gradients were always prepared at least 24 h prior to the experiment and were kept at 4°C until required. Five ml of infected tissue homogenate were deposited gently on top of the gradient avoiding mixing with the topmost layer. The gradients were centrifuged at the specified speed and time using a refrigerated benchtop centrifuge with a swingout rotor. Tubes were removed carefully from the centrifuge, and the phases were carefully and independently collected by using a 1 ml pipette and wide-bore tips.
The recovered phases were each diluted ten times with 1x IB and centrifuged at 1080 g for 15 min at 4°C. The resulting pellets were resuspended in cold 1x IB and used for further experimentation.

4.2.1.3. General steps for the preparation of self-formed continuous Percoll gradients

To make self-formed continuous Percoll gradients, a single Percoll-tissue homogenate suspension was prepared to a total volume of 80 ml from 100% Percoll, 10x IB, sterile MilliQ water and homogenate of infected tissue in proportions adjusted to obtain the desired final Percoll concentration (20-30% v/v) and an isolation buffer concentration of 1x. This solution was split into four 45 ml polycarbonate round-bottom tubes and centrifuged at 25,000 g for 30 min in a fixed angle rotor (Sorvall SS34) at 4°C without braking. Tubes were removed carefully from the centrifuge, and fractions of the desired volume were carefully and independently collected with a 1 ml pipette and wide bore tips. The recovered fractions were diluted independently ten times with 1x IB and centrifuged at 1080 g for 15 min at 4°C, using a centrifuge with a swingout rotor. The resulting pellets were resuspended in cold 1x IB and used for further experimentation.

4.2.2. Purification of haustoria from enriched fractions by fluorescent-activated cell sorting (FACS)

4.2.2.1. Concanavalin A- Alexa 488 staining

The stock solution of Concanavalin A, Alexa Fluor 488 Conjugate (Molecular Probes)(ConA-488) was prepared at a final concentration of 1 mg/ml as per the manufacturer’s instructions, but without the addition of sodium azide. The stock solution was stored in 100 µl aliquots at -20°C in aluminium foil to protect them from light exposure. Enriched suspensions of haustoria obtained from Percoll gradients were diluted ten times with 1x IB and centrifuged at 1080 g for 15 min at 4°C, using a centrifuge with a swingout rotor. The resulting pellets were resuspended in 4 ml of 1x IB (referred to as the haustorial suspension) and transferred to a glass vial with plastic lid, 6 ml capacity. Two 100 µl aliquots of ConA-488 were added to the haustorial suspension avoiding light exposure at all times, and the glass vial was mixed gently on a rotary mixer for 20 min at room temperature or for 45 min at 4°C. Subsequently, the haustorial suspension was centrifuged at 4,000 g in a benchtop centrifuge for 5 min at 4°C and the pellets were washed twice with cold 1x IB. The resulting pellets were resuspended in 4 ml of 1x IB and kept on ice until sorting.

4.2.2.2. Fluorescent-Activated Cell Sorting of ConA-488 stained haustoria
Chapter 4: Purification of Pst haustoria

Haustorial suspensions stained with ConA-488 were diluted 1:10 with chilled and sterile 1x phosphate buffered saline (1x PBS: 8 g NaCl, 0.2 g KCl, 1.44 g Na₂HPO₄, 0.24 g KH₂PO₄, dissolved in 1 L of MilliQ water, final pH 7.4) to adjust the concentration of the enriched haustorial preparation containing both chloroplasts and haustoria to approximately 2.5–5.0x10⁶ particles/ml. A Dickinson BD FACSARIA II cell sorter (BD Biosciences) was set up with parameters optimized for separation of fungal haustoria such as 100 µM nozzle and 20 psi sheath pressure. Buffer 1x PBS was used as the sheath fluid and cells were run using the lowest possible sample pressure with a resulting sample rate of 5,000-10,000 events per second. A 13 mW 488 nm solid state laser (Sapphire, Coherent Inc. Santa Clara, CA) was used to excite the Alexa 488 fluor and fluorescence was detected using a 502 nM longpass and 530/30 nM bandpass filter. Chlorophyll autofluorescence was excited with the same 488 nm laser and detected using a 655 longpass and 695/40 bandpass filter. Forward and side scatter information was used to identify and exclude doublets by comparing the forward scatter height and width and the side scatter height and width. To optimise the forward scatter signal, a neutral density 2 filter was used in front of the forward scatter detector. Cells were collected in 5 ml glass tubes containing 1x PBS, keeping them at 4°C and agitating them periodically at 300 rpm to prevent settling. Samples obtained after FACS purification were pooled and centrifuged at 1080 g for 15 min at 4°C. The supernatant was discarded and the final pellet was snap-frozen in liquid nitrogen, and stored at −80°C until required. All FACS experiments were performed in collaboration with Dr Harpreet Vohra and Mr Michael Devoy from the Microscopy & Cytometry Resource Facility at the John Curtin School of Medical Research, The Australian National University.

4.2.3. Viability assays on sorted haustoria

4.2.3.1. 7-amino actinomycin D staining of dead cells

Haustorial suspensions were concentrated by centrifugation at 1080 g for 15 min and resuspended at 5 x 10⁵ particles/ml in 1x PBS. Fractions of 100 µl were incubated for 10 min with 2 µl 7-amino-actinomycin D (7-AAD, stock solution 0.1 mg/ml; Sigma) at room temperature and in darkness. Samples were directly evaluated by microscopy and flow cytometry. 7-AAD fluorescence was exited at 488 nm (250 mW). Positive controls for this test were made by treating samples with 2 µl of concentrated extran MA 03 (Merk Millipore) for 5 min to induce damage on haustoria plasma membranes, prior to incubating the samples with 7-AAD. Microscopy was conducted using a Leica SP5 inverted confocal microscope and HCX PL APO lambda blue 63.0x1.40 OIL UV.
4.2.3.2. CellTracker Orange CMRA viability test

Haustorial suspensions were concentrated by centrifugation at 1080 g for 15 min and resuspended at 5 x 10^5 particles/ml in 1x PBS. Fractions of 100 µl were incubated for 30 min with 0.6-5 µM CellTracker Orange CMRA: 9′-(4 [and 5]-chloromethyl-2-carboxyphenyl)-7′-chloro-6′-oxo-1,2,2,4-tetramethyl-1,2-dihydropyrido[2′,3′-6]xanthene (Molecular Probes) in darkness. Treated samples were centrifuged at 1080 g for 15 min, resuspended in 1 ml of 1x PBS and incubated for 30 min in darkness. Samples were pelleted and resuspended in 10 µl for microscopic analyses. CMRA fluorescence was excited at 561 nm with a laser emitting at a maximum wavelength of 576 nm. Negative controls for this test were made by treating samples with formaldehyde (4% w/v for 30 min), heat (65°C for 10 min) or a prolonged time at room temperature (5 days), prior to applying the CellTracker Orange CMRA viability test. Microscopy was conducted using a Leica SP5 inverted confocal microscope and HCX PL APO lambda blue 63.0x1.40 OIL UV.

4.3. RESULTS

4.3.1. Enrichment of haustoria by discontinuous Percoll gradients

After homogenization of the infected wheat leaves and filtration to remove plant cell debris, I tested if haustoria could be enriched from the homogenate on the basis of density using isopycnic centrifugation. Percoll was chosen as the medium for these preparations because it has been shown to be useful in a wide variety of biological samples, has low toxicity towards cells, and can form both continuous and discontinuous gradients. The osmolality of Percoll was adjusted by adding 1 part of 10x IB to 9 parts of concentrated Percoll, generating a stock solution of 90% Percoll. From this stock solution, subsequent dilutions to the desired concentration of Percoll were made with 1x IB. Without prior knowledge of the buoyant density of Pst haustoria on Percoll gradients, discontinuous gradients were used first to determine their isopycnic point (that point where their buoyant density is equal to that of the gradient). Multi-step gradients covering a wide range of densities were prepared by layering 5 ml of Percoll solutions of different concentrations (20-70% v/v) in increments of 10%, starting with the most dense at the bottom of the tube to the least dense on the top (Figure 4.2). The preparation of the gradients was most conveniently done 24 h prior to use. It was important to press the pipette tip against the wall of the tube just above the surface of the liquid while depositing each new layer to avoid splashes that could cause mixing of the interfaces. A clear
differentiation between the gradient steps was visible by inspection. Five ml of the homogenate of infected tissue was loaded on top of each gradient and the tubes were centrifuged as shown in Figure 4.2.

![Figure 4.2. Distribution of *Pst* haustoria on a discontinuous Percoll gradient. Five ml of homogenate of infected tissue was layered onto a Percoll multi-step gradient and centrifuged in a swing-out rotor. After centrifugation, fractions were recovered and haustoria were counted using a haemocytometer. The numbers of haustoria were plotted as percentages relative to the total number of haustoria contained in the whole gradient. The bar graph represents the results obtained from three biological replicates of the experiment.]

Gentle centrifugation conditions were used as low speed and short centrifugation run time are critical for the isopycnic banding of cells at the proper interfaces on discontinuous Percoll gradients (Pertoft et al., 1978, Amersham-Biosciences, 2001). After centrifugation, each of the phases was independently recovered and diluted 1:10 with 1x IB to reduce the Percoll concentration and improve the pelleting of haustoria by centrifugation. The distribution of haustoria in each fraction was evaluated by counting them in each of the phases, and the results were plotted as percentages as shown in Figure 4.2. The haustoria were distributed most heavily (~82%) in the 20% and 30% phases, suggesting that these are close to the buoyant density of *Pst* haustoria.
Based on the results obtained from the multi-step gradients, a simpler three-step discontinuous gradient was used to assess the level of chloroplast contamination in those phases where most haustoria sedimented (Figure 4.3). The volume of each step was increased so that the particles in the homogenate of infected tissue would travel a similar distance during centrifugation as in the previous experiment, so that the particles located at the interfaces would be better separated.

**Figure 4.3.** Measurement of the haustoria:chloroplast ratio in a three-step discontinuous Percoll gradient. **A.** Five ml of homogenate of infected tissue was layered onto a three-step Percoll gradient and centrifuged in a swing-out rotor. The dashed lines indicate the distinguishable interfaces after centrifugation, and the green colouration indicates the areas of chloroplast concentration before and after centrifugation. **B.** Bright field microscopic images of an aliquot of each phase after centrifugation. Haustoria and chloroplast numbers in each phase were counted using a haemocytometer, and the haustoria:chloroplast ratios are indicated above each image. The black arrows indicate haustoria. Bright field images collected on a Leica DMR epifluorescence microscope, 20x magnification.

After centrifugation, four phases were clearly distinguishable (Figure 4.3). The 20% and 30% steps of Percoll where haustoria were expected to migrate remained constant in volume after centrifugation and were numbered as phases 3 and 4 respectively. Phases 1-4, the interface between phases 2 and 3 (~1 ml), and the pellet were collected independently for analysis of their contents. Viewing of the fractions by bright-field microscopy revealed that plant cell debris and broken fungal hyphae mostly sedimented in the pellet, although they were also found at low levels in phases 1-4. Conversely,
chloroplasts were visibly concentrated at the interface between phases 2 and 3 and in the pellet, but were the major source of contamination in all phases. Haustoria and chloroplasts were counted in samples from phases 1-4 by haemocytometer. The interface of phases 2 and 3 and the pellet were not included in this analysis as their very high levels of chloroplast contamination made them unsuitable for further experimentation. The percentage of haustoria in all phases were calculated in relation to the total number of haustoria in each phases, and the level of chloroplast contamination expressed as the haustoria:chloroplast ratio. This ratio is used as a standard point of comparison with the ConA-affinity column method. Remarkably, nearly $6.8 \times 10^6$ haustoria were recovered from phases 1-4 when starting from ~15 g of heavily infected tissue, mostly concentrated in phases 3 (~55%) and 4 (~36%), where the haustoria:chloroplast ratio was ~1:20 and ~1:50 respectively. This indicated that although the bulk of haustoria were recovered from phase 3 (30% Percoll), these haustoria are two times more contaminated with chloroplasts than those in phase 4. This experiment suggested that *Pst* haustoria could be enriched from plant cell debris, infectious fungal hyphae and chloroplasts on the basis of density and sedimentation on a discontinuous Percoll gradient.

In isopycnic centrifugation, cells of a particular density sediment until they reach the point at which their density is the same as that of the gradient (Patel, 2000). Differences in cell size merely affect the rates at which cells reach their equilibrium positions (Patel, 2000). The speed and time of centrifugation on discontinuous Percoll gradients can have a profound effect on the migration and final distribution of the particles contained in different biological samples, especially if the equilibrium point is unknown (Makler et al., 1998, Zalups and Lash, 1996). Thus, similar three-step gradients were prepared and shorter centrifugation times were tested to investigate the effect on haustorial yield, distribution and chloroplast contamination across the phases (Table 4.1.).

<table>
<thead>
<tr>
<th>Phases analysed</th>
<th>Centrifugation time</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10 min</td>
<td>15 min</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Haustoria distribution</td>
<td>Haustoria:chloroplasts ratio</td>
<td>Haustoria distribution</td>
</tr>
<tr>
<td>2</td>
<td>8.3%</td>
<td>1:17</td>
<td>4.6%</td>
</tr>
<tr>
<td>3</td>
<td>65%</td>
<td>1:34</td>
<td>56%</td>
</tr>
<tr>
<td>4</td>
<td>26.7%</td>
<td>1:40</td>
<td>39.4%</td>
</tr>
</tbody>
</table>

Table 4.1. Effect of centrifugation time on the distribution and purity of haustoria on a three step Percoll gradient. Five ml of homogenate of infected tissue was layered onto a three-step Percoll gradient prepared as indicated in Figure 4.2. Gradients were centrifuged for 10 or 15 min in a swing-out rotor at 720 g. Haustoria and chloroplasts were counted in the
resulting phases (except for the top phase which was devoid of haustoria) and the numbers were converted to haustoria percentages and haustoria:chloroplast ratios. The asterisks indicate that the ratio numbers are only indicative as they can vary slightly from one experiment to another depending on the infection density.

Reducing the centrifugation time increased the retention of haustoria and chloroplast in the less dense phases. A reduction from 20 min to 10 min centrifugation raised the percentage of haustoria in phase 3 from 55% to 65%, although the ratio of contamination also increased from ~1:20 to ~1:34. There were few differences between centrifugation for 15 or 20 min, suggesting that haustoria reach their equilibrium point at or shortly after 15 min on the three-step Percoll gradients described here.

Taking into account the variables assessed previously, a final step gradient was designed and tested to optimize haustoria recovery in the 20% phase and minimize chloroplast contamination. In this experiment, the lowest step of Percoll was set at 40% and was smaller in volume (5 ml) to create a density barrier for the haustoria that normally would sediment in 30% Percoll. The 20% step was larger in volume (12 ml) to compensate for the reduction in volume of the bottom step, and thus maintain the total migration distance. The uppermost step was the same as in previous experiments, but the time of centrifugation was reduced to 12 minutes. The scheme is illustrated in Figure 4.4.

**Figure 4.4. A three-step Percoll gradient for haustoria enrichment.** Five ml of homogenate of infected tissue was layered onto a three-step Percoll gradient and centrifuged in a swing-out rotor. The concentration and original volume of each Percoll step are indicated. The resulting phases and their volumes after centrifugation are indicated. All phases were recovered and phase 4 was treated with ConA-488 to specifically stain haustoria (green fluorescence). Fluorescent images were collected on a Leica DMR epifluorescence microscope, 20X magnification.
After centrifugation, six phases could be clearly differentiated as well as a pellet at the bottom of the tube (Figure 4.4). Each phase was removed carefully and diluted 1:10 with 1x IB to reduce the Percoll concentration and facilitate pelleting of the particles by centrifugation. Close to 75% of haustoria were retained in phase 4 with a haustoria:chloroplast ratio of ~1:38. More than $5.2 \times 10^6$ haustoria were present in phase 4 starting from 15 g of heavily infected tissue. Although the level of contamination was increased ~2.5 times when compared to the ConA affinity column method (discussed in Chapter 2), the yield of this discontinuous Percoll gradient was increased ~8.6 times relative to the column procedure. This result showed that the yield of Pst isolated haustoria was improved remarkably by the use of discontinuous Percoll gradients, and set the conditions to use this technique as an alternative to affinity columns.

**4.3.2. Enrichment of haustoria using self-formed continuous Percoll gradients**

As an alternative to discontinuous gradients, the centrifugation of Percoll solutions at $>10,000$ g in 0.15 M NaCl or at 25,000 g in 0.25 M sucrose, for at least 15 min in fixed angle rotors results in a self-formed density gradient (Vincent and Nadeau, 1984, Amersham-Biosciences, 2001). The size heterogeneity of the colloidal silica particles of Percoll results in their uneven distribution during centrifugation, forming a density gradient which becomes progressively steeper with g-force and centrifugation time. Percoll can thus be mixed with the suspension of biological particles (cells or subcellular particles) prior to centrifugation, so that the particles will band isopycnically as the continuous gradient is formed *in situ*. Thus, these gradients were also explored as an option to enrich haustoria from homogenates of infected tissue.

Guided by the literature on the characteristics of self-formed continuous Percoll gradients (Amersham-Biosciences, 2001), and the information obtained above on the distribution of haustoria on discontinuous Percoll gradients, *in situ* density gradients from solutions of 20% and 30% Percoll (v/v) were tested (from here on; P20 and P30). The solutions were made by mixing the homogenate of infected tissue, with 10x IB to a final concentration of 1x, and Percoll to final concentrations of 20% or 30% respectively. The conditions of centrifugation were 25,000 g for 30 min at 4°C. After centrifugation, fractions of 1 ml were collected from the top to the bottom of the gradients, and haustoria were counted in each fraction (Figure 4.5).
Figure 4.5. Self-formed continuous Percoll gradient for haustoria enrichment. A. Premixed solutions of homogenate of infected tissue and Percoll to the indicated concentrations were centrifuged at 25,000 g for 30 min and the resulting self-formed gradients are pictured. The white dashed lines represent the levels corresponding to fractions 12 and 10 in the 20% and 30% gradients respectively B. Haustoria were counted in samples from each 1 ml fraction recovered from the top to the bottom of the gradient. The distribution of haustoria across the gradient is shown as percentages.

In P30, the haustoria were located mostly in the top half of the gradient, as nearly 90% of the haustoria could be recovered in the first ten fractions. In P20, haustoria were spread more evenly along the gradient, only ~47% were recovered in the first 10 fractions and approximately 20% of the haustoria sedimented in the pellet. The largest percentage of haustoria recovered in a single fraction was ~35% in fraction 2 of P30. However, this fraction was very close to one of the two major areas of concentration of chloroplasts in P30, as shown in Figure 4.5A. Plant cell debris and broken hyphae were found mainly in the pellets of both gradients and at minor levels throughout the gradients. On the other hand, the chloroplasts were mainly concentrated after fraction 14 in P20, and between phases 3-6 and 12-17 in P30, although none of the fractions in either gradient were free of chloroplast contamination. These results show that neither of these gradients successfully sedimented a high percentage of haustoria in a single fraction separate from chloroplasts, similar to the discontinuous gradients.

The distribution of haustoria and apparent distribution of chloroplasts in P20 and P30 suggested that collecting fractions 1-12 from P20, and 1-10 from P30, would be a good strategy to maximize haustoria recovery. To investigate the levels of chloroplast contamination after pooling those fractions, the haustoria:chloroplast ratio was calculated by haemocytometer. Approximately 56% of
haustoria were recovered with a haustoria:chloroplast ratio of ~1:17 from P20, whereas ~90% of haustoria were recovered from P30 but with about four times more chloroplast contamination, as the haustoria:chloroplast ratio was ~1:60. These results suggested that the concentration of Percoll affected the distribution of haustoria and contaminants throughout the resulting gradients. In none of the cases were haustoria located in a single fraction, although a higher concentration of Percoll seemed to compress haustoria into a narrower region of the gradient.

Given that haustoria appear to be confined to a more narrow level of the gradient in P30, I undertook additional approaches to determine their banding media density. A P30 gradient was prepared and an additional treatment replacing the homogenate of infected tissue with 1x IB was included. Fractions of 1 ml were recovered from the buffer-only gradient and the density was estimated by dividing the weight of each fraction in mg by their volume (1 ml). Haustoria were counted from fractions recovered from the gradients prepared with infected material, and the percentages in each fraction were calculated. The distribution of haustoria and the corresponding fraction density were plotted together in Figure 4.6.
Chapter 4: Purification of Pst haustoria

Haustoria were mostly concentrated in two areas of the gradient with densities between ~1,028-1,043 mg/ml and ~1,061-1,064 mg/ml. Within these two regions, no differences in haustoria size or cytoplasmic appearance (granularity) were detected by bright-field microscopy. However, differences in age, developmental stage and internal/external complexity cannot be ruled out as extensive microscopic or metabolic analyses were not done. Collection of the first ten fractions allowed the recovery of up to ~6.1x10^6 haustoria from 15 g of heavily infected tissue, and the haustoria:chloroplast ratio varied from 1:55 to 1:76. Compared with the ConA-affinity method, the self-formed continuous 30% Percoll gradient allowed a ~10-fold increase in the haustoria recovery efficiency, although these cells were ~5 times more contaminated with chloroplasts. The main methods designed and tested in this chapter (except the ConA affinity column) to enrich haustoria from Pst-wheat infected tissue are summarized in Table 4.2.

<table>
<thead>
<tr>
<th>Isolation method</th>
<th>Phase/fraction collected</th>
<th>Haustoria recovered</th>
<th>Efficiency</th>
<th>Haustoria chloroplast ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>ConA affinity columns</td>
<td>-</td>
<td>6.0 x 10^5</td>
<td>1</td>
<td>1:15</td>
</tr>
<tr>
<td>3-Step gradient (10% -20%-30%)</td>
<td>Phase 3 (20% Percoll)</td>
<td>3.74 x 10^6</td>
<td>6.2</td>
<td>1:20</td>
</tr>
<tr>
<td>3-Step gradient (10%-20%-40%)</td>
<td>Phase 4 (20% Percoll)</td>
<td>5.2 x 10^6</td>
<td>8.6</td>
<td>1:38</td>
</tr>
<tr>
<td>Self-formed continuous gradient P20</td>
<td>Fractions 1-12</td>
<td>3.79 x 10^6</td>
<td>6.3</td>
<td>1:17</td>
</tr>
<tr>
<td>Self-formed continuous gradient P30</td>
<td>Fractions 1-10</td>
<td>6.1 x 10^6</td>
<td>10</td>
<td>1:60</td>
</tr>
</tbody>
</table>

Table 4.2. Comparison of the haustoria isolation methods applied on Pst infected tissue. The total number of haustoria isolated from 15 g of heavily infected tissue using the ConA affinity method was used as the reference to calculate the efficiency of Percoll gradients (i.e. -fold times the number of isolated haustoria) from the same amount of infected tissue. The asterisks indicate that these values are only indicative as they can vary slightly from one experiment to another depending on the infection density.

There was a significant improvement in the haustoria yield in all cases when applying Percoll gradients to isolate haustoria from infected tissue, compared to the ConA affinity method. Both
continuous and self-formed continuous gradients could be modified to achieve higher yields, albeit with higher levels of chloroplast contamination. Discontinuous gradients allowed a higher level of purity than self-formed continuous gradients, with only small differences in yield, although their preparation required a greater degree of technical proficiency. The consecutive use of continuous and discontinuous gradients might be an alternative to reduce the level of chloroplast contamination while maintaining a higher haustorial yield than that obtained with the ConA affinity method. The time required to complete the whole process for isolating Pst haustoria by Percoll gradients was comparable to that required for ConA affinity columns (~3.5 h), another favorable point for the use of these gradients as a routine method to isolate haustoria.

4.3.3. Haustoria purification by flow cytometry

Haustoria isolated by ConA affinity columns or Percoll gradients were analyzed routinely by fluorescent microscopy using lectins conjugated to fluorescent tags (ConA-Alexa Fluor 594 and WGA-FITC). ConA-Alexa 594 exhibited consistent and very high specificity for haustoria, but no or very low affinity for contaminants such as broken hyphae or plant cell debris, and no affinity for chloroplasts. To investigate if the fluorescent labeling of haustoria could be exploited to further separate these cells from chloroplasts, fluorescence-activated cell sorting (FACS) was explored. The excitation and emission wavelengths govern the possible combination of fluorophores for flow cytometry. The use of multiple fluorophores with similar excitation wavelengths and different emission wavelengths (or “colors”) makes it possible to differentiate cell types or subpopulations by measuring their signals with a single laser in a flow cytometer. Thus, choosing the correct combination of fluorophores was critical for successful separation of haustoria and chloroplasts. Initial tests showed that chloroplasts could be excited with a 488 nm laser and detected using a 655 longpass and 695/40 bandpass filter (red channel). However, ConA-Alexa 594 and chlorophyll were a poor combination as their emission spectra were too close. Thus, the commercially available Concanavalin A, Alexa Fluor 488 conjugate (ConA-488) was used to label haustoria (green channel). This fluorophore was excited with a 488 nm laser and its emission was detected using a 502 nM longpass as well as a 530/30 nM bandpass filter, well separated from the chlorophyll emission wavelength.

Enriched suspensions of haustoria obtained by self-formed continuous 30% Percoll gradients (from here on called pre-sorted samples) were treated with ConA-488 and washed to remove unbound fluorescent lectin before FACS. To set fluorescence and scatter parameters, chlorophyll-containing
particles were identified and excluded based on their chlorophyll autofluorescence. ConA-488 positive particles were identified and selected for collection (Figure 4.7). Forward and side scatter information was used to identify and exclude doublets (when more than one particle is within the droplet interrogated by the laser beam) by comparing the forward scatter height and width with the side scatter height and width.

The data show that FACS was successful in discriminating haustoria from chloroplasts based on their fluorescence patterns (Figure 4.7). While chloroplasts had high-red low-green fluorescence,
haustoria had high-green low-red fluorescence. The sorting process takes on average two hours, although this can vary with the concentration of the samples and the sorting speed, which in these experiments was kept low as the effect of high pressure on isolated haustoria cells is unknown. Microscopic analyses of pre- and post-sorted samples showed that sorted populations of haustoria contained very low levels of contamination (Figure 4.7C-D) and appeared to be intact after sorting. Quantitative purity tests were done by re-sorting populations of sorted haustoria, resulting in haustoria samples with 95-98% purity (Figure 4.7B). On average, the haustoria:chloroplast ratio in post-sorted haustoria samples was 49:1. Between 18-25% of the original number of haustoria in the enriched fractions from Percoll gradients were recovered after sorting, demonstrating the efficiency of this technique.

4.3.4. Viability tests on sorted haustoria

4.3.4.1. Haustoria viability tested by 7-amino actinomycin D (7-AAD) dye exclusion

One method to distinguish between dead and live cells in a population is by dye exclusion. Live cells have intact membranes that exclude a variety of dyes that easily penetrate the damaged, permeable membranes of non-viable cells. Among the different fluorochromes that can be used to stain non-viable cells, 7-amino actinomycin D (7-AAD) is one of the most widely used viability assays for microscopic and flow cytometric analyses (Schmid et al., 1992). 7-AAD is a membrane impermeant dye that is generally excluded from viable cells. It binds to double stranded DNA by intercalating between base pairs in G-C-rich regions. 7-AAD can be excited at 488 nm with a laser, emitting at a maximum wavelength of 647 nm, which can be detected using a 655 nm long-pass filter, similarly to chlorophyll (red channel).

Following the manufacturer’s instructions, 7-AAD was tested on pre-sorted haustoria samples stained with ConA-488. Samples were incubated at room temperature for 10 min with 7-AAD at a final concentration of 2 µg/ml, and were analyzed immediately using confocal scanning laser microscopy (CSLM) (Figure 4.8).
Haustoria and chloroplasts were clearly differentiated by their fluorescence patterns; green haustoria (Figure 4.8C-D) and red chloroplasts (Figure 4.8B-D). If haustoria were permeable by 7-AAD they would have become internally fluorescent and would be stained red in panels B and D of Figure 4.8. However, they could not be detected in the red channel even 1 h after the dye was added.

To verify the role of membrane impermeability in haustorial resistance to dye uptake, an aliquot of the sample was incubated for 5 min with 2% (v/v) of the detergent extran MA 03 to increase the permeability of the haustoria membranes. The sample was incubated subsequently with 7-AAD for 10 min and analyzed immediately using CSLM (Figure 4.9).
Chapter 4: Purification of Pst haustoria

Figure 4.9. Pre-sorted haustoria treated with extran and then 7-AAD. The sample was incubated with extran for 5 min, then with 7-AAD for 10 min before visualisation by CSLM exciting at 488 nm. Single focal plane images were taken in bright field (A), red channel (B; chloroplasts and 7-AAD-stained haustoria), green channel (C; ConA-488 stained haustoria) and merged red and green channels (D). The white rectangles in panel B indicate the positions of haustoria in the sample. Fluorescent images were collected from a single focal plane on a Leica SP5 inverted confocal microscope.

The integrity of haustoria appeared not to be affected by the extran treatment as the cells were intact (Figure 4.9C). Haustoria fluoresced strongly in the red channel throughout the area demarcated by the haustorial cell wall (Figure 4.9B and D), indicating that their membranes became dye-permeable after extran treatment. This assay was useful to adjust the conditions and treatments to used 7-AAD on Pst haustoria and is consistent with the idea that haustoria remain alive after enrichment on continuous Percoll gradients.

Next, haustoria were sorted in aliquots of ~1x10^5 cells based on ConA-488 labelling (green fluorescence). Half of the aliquots were treated with extran, and then all aliquots were treated with 7-AAD. Preliminary microscopic analyses showed that sorted haustoria treated with 7-AAD were impermeable to the dye. Conversely, sorted haustoria treated with extran and 7-AAD showed consistently higher permeability (Figure 4.10).
Figure 4.10. Microscopic analysis of pure haustoria treated with 7-AAD. Images of single haustoria in bright field, green channel (Con-488), red channel (7-AAD) and merged red and green channels after treatment as indicated on the left of the images. Fluorescent images were collected on a single focal plane on a Leica SP5 inverted confocal microscope.

The same samples were further analyzed by flow cytometry to determine the percentages of live and dead haustoria after sorting. Flow cytometric analyses of aliquots of sorted haustoria treated with 7-AAD and extran + 7-AAD are shown in Figure 4.11.
Figure 4.11. FACS analysis of pure haustoria tested for viability with 7-AAD. A. Sorted haustoria treated with 7-AAD and re-sorted. B. Sorted haustoria treated with extran, then with 7-AAD, and re-sorted.

Haustoria treated with extran showed a much higher level of red fluorescence compared to untreated haustoria, becoming a shifted undefined population in the scatter diagram. This increase in red fluorescence suggests that the treatment with extran makes the haustoria cells permeable to the 7-AAD stain, probably by damaging their membranes. Approximately 5% of these cells were discriminated as nonviable cells using 7-AAD on sorted haustoria (not treated with extran), indicating that approximately 95% of the sorted haustoria remained alive after the sorting process.

4.3.4.2. Haustoria viability tested using CellTracker CMRA

A second viability test was applied on sorted haustoria using an alternative principle to dye exclusion. CellTrack fluorescent probes (Invitrogen) are a group of reagents that pass freely through cell membranes, but once inside the cell, are transformed into cell-impermeant reaction products. The cells where this occurs must be viable for the system to work. The dye CMRA belongs to the CellTracker group and was used to assess haustoria viability. CMRA can be excited at 548 nm with a laser light, emitting at a maximum wavelength of 576 nm. To evaluate the utility of CMRA viability test and to standardize the conditions to apply it on haustoria, pre-sorted haustoria stained with ConA-488 were treated with different concentrations of CMRA following the manufacturer’s instructions. Aliquots of the same original sample were adjusted to contain ~5x10^5 particles before the treatments. Although this dye can be used at a final concentration of up to 25 µM, it is
recommended to keep the concentration as low as possible to maintain normal cellular physiology and reduce potential artifacts. Thus, the samples were incubated with 0.6 µM, 1.25 µM, 2.5 µM or 5.0 µM CMRA for 30 min, and after pelleting the cells and removing the supernatant containing the dye, the cells were incubated for 30 min in 1X PBS, always in darkness. Cells were pelleted, resuspended in 1x PBS, and analyzed by CSLM to determine the optimal CMRA concentration and microscopy conditions to assess haustorial viability (Figure 4.12).

Figure 4.12. Standardization of CMRA concentration to assess haustorial viability in pre-sorted samples. Aliquots from enriched suspensions of haustoria were treated with different concentrations of CMRA following the manufacturer’s instructions. The treated samples were analyzed by CSLM. The concentration of the dye is indicated on the upper left corner of each image. The strong blue fluorescence obtained with 5.0 µM CMRA corresponded with the location of haustoria. Fluorescent images were collected on a single focal plane on a Leica SP5 inverted confocal microscope.
The microscope was configured to enable high sensitivity to CMRA fluorescence and reduced background noise during acquisition. Untreated samples were used as negative controls to set the threshold of fluorescence detection. Haustoria in samples treated with 0.6 µM CMRA did not show any fluorescence. Haustoria in samples treated with CMRA 1.25 µM and 2.5 µM showed very low levels of fluorescence that difficult to distinguish from the background. Conversely, haustorial samples treated with CMRA 5.0 µM exhibited higher levels of fluorescence (blue color) making them easily distinguishable in the observed fields (Figure 4.12).

To investigate if the fluorescence observed in haustoria was due to their active metabolic status and not an artifact of the treatment, additional samples were first subjected to treatments to render haustoria unviable prior to treatment with CMRA 5.0 µM (Figure 4.13). These treatments included; incubation in 4% formaldehyde (w/v) for 30 min, incubation at 65°C for 10 min, or incubation in 70% isopropanol for 30 min.
Figure 4.13. CMRA viability test on pre-sorted haustoria samples previously subjected to heat and chemical treatments. Incubation in 4% formaldehyde for 30 min (left panels) or at 65°C for 10 min (central panels) on two pre-sorted samples. An untreated sample was used as a positive control (right panels). Subsequently all the samples were treated with 5.0 µM CMRA and were analyzed by CSLM (green channel for ConA-488 and blue channel for CMRA). Fluorescent and bright field images were collected on a Leica SP5 inverted confocal microscope.

Samples subjected to formaldehyde and high temperature treatments showed a significant reduction in CMRA fluorescence when compared to untreated samples. However, there was a remnant level of background fluorescence in both treatments (Figure 4.13). Closer examination of individual haustoria indicated that despite this background, haustoria from untreated samples were
easily discriminated from those subjected to high temperature or formaldehyde (Figure 4.13B). This suggests that the CMRA viability tests can discriminate between viable and non-viable haustoria and can therefore be useful to assess the viability of post-sorted haustoria.

All the particles in samples treated with isopropanol became massively and nonspecifically fluorescent in both the green and blue channels, indicating that this treatment is not compatible with the viability assay (Figure 4.14).

![Figure 4.14. CMRA viability test applied to haustoria-enriched samples treated previously with 70% isopropanol.](image)

Although heat and formaldehyde treatments can be used as negative controls for the CMRA viability test, the background fluorescence exhibited by haustoria in those treatments led me to explore a further treatment to generate a better negative control. My previous observations suggested that haustoria isolated using the ConA-affinity method or on Percoll gradients preserve their shape and internal complexity for up to one week when maintained at 4°C (data not shown). It seems feasible that after such a long period out of their natural environment suspended in a media devoid of any nutrient, the physiological and metabolic status of haustoria would be compromised, and as such provides a good negative control for this assay. Thus, freshly sorted haustoria (0 days post sorting, dps) and sorted haustoria kept at room temperature for 5 days (5 dps) were treated with 5.0 µM CMRA to compare their viability (Figure 4.15).
The bright field images revealed that haustoria at 0 and 5 dps had identical appearances, retaining their granular cytoplasmatic contents and with intact membranes (Figure 4.15, left panels). However, they responded completely differently in the blue channel, where 0 dps haustoria exhibited significantly more internal fluorescence than 5 dps haustoria. Although the blue fluorescence in 0 dps haustoria was not distributed evenly across the cytoplasm, most of the haustoria observed had a consistently high level of fluorescence, contrary to what was observed in 5 dps haustoria (Figure 4.15A). The CMRA viability test on sorted haustoria suggested that these cells remain alive after the sorting process, making them amenable for further experimentation.

4.4. DISCUSSION

The haustorium is the most important feature of biotrophic fungi (Mendgen and Hahn, 2002). Taxonomically diverse groups of obligate plant pathogens, namely the rust fungi (Basidiomycetes), the powdery mildew fungi (Ascomycetes), and the downy mildew (Oomycetes), have independently evolved haustoria, suggesting that it represents a successful adaptation for growth within living plant tissue (Hahn et al., 1997b, Catanzariti et al., 2011). Despite the importance of the haustorium in biotrophy, they have been difficult to study due to their close association with host cells and the inability to manipulate fungi to produce haustoria in vitro. In this study, the isolation of Pst haustoria from infected wheat tissue has been a fundamental strategy for the transcriptomic analysis of this pathogen. In previous Chapters, I described how Pst haustoria were isolated using the ConA affinity chromatography method developed by Hahn and Mengden in 1992. However, the very low yield of haustoria per gram of infected tissue using this method was a major limiting factor in generating sufficient material for high-throughput sequencing. Additionally, the large number of contaminating chloroplasts that co-purified with haustoria would prevent the development of the proteomics studies planned as part of this project. These disadvantages of the ConA-affinity technique led me to explore Percoll density gradients to improve the efficiency of recovery of haustoria, combined with FACS to improve their purity.

4.4.1. Haustoria isolation by discontinuous and continuous Percoll gradients

Figure 4.15. CMRA viability test applied on sorted haustoria. Haustoria 0 and 5 dps were treated with 5.0 µM CMRA and were analyzed by CSLM (green channel for ConA-488 and blue channel for CMRA). Fluorescent and bright field images were collected on single focal plane on a Leica SPS inverted confocal microscope.
To separate *Pst* haustoria from other fungal structures and contaminating plant components I used isopycnic centrifugation on Percoll gradients, used previously by others to isolate haustoria and intracellular hyphae of other fungi (Pain et al., 1994, Micali et al., 2011, Takahara et al., 2009). Percoll is a commercially available medium with consistent physicochemical properties between batches, giving reproducibility over the course of an experiment. Three different strategies can be used to form Percoll gradients: step density gradients (discontinuous gradients); self-formed gradients (continuous gradients) and preformed continuous gradients. Step density gradients often give good separation of cells, but are not as sensitive or selective as continuous gradients (Pertoft, 2000). Thus, with no previous data on the use of Percoll gradients for isolation of *Pst* haustoria, I investigated discontinuous and continuous Percoll gradients.

Discontinuous gradients were initially tested to get a general feeling for the Percoll concentration which would trap the majority of *Pst* haustoria. A multiple-step gradient covering a wide range of Percoll concentrations showed that haustoria were mostly retained in the phase corresponding to 20% Percoll, and secondarily in the phase corresponding to 30% Percoll. Based on this result, a simpler three step Percoll gradient (10%, 20% and 30%) was designed to investigate how chloroplasts and other contaminants would sediment with respect to haustoria. Fungal structures other than haustoria sedimented in the pellet, along with plant cell debris and a proportion of the chloroplasts, enormously facilitating the enrichment of haustoria. However, large numbers of chloroplasts were distributed in all phases of the gradient. This distribution agrees to a certain extent to what has been reported for wheat and spinach chloroplasts separated by Percoll step gradients. Whereas intact chloroplasts sediment at Percoll concentrations higher than 40%, broken or damaged chloroplasts sediment at the less concentrated Percoll concentrations (Seigneurin-Berny et al., 2008, Kamal et al., 2012). Nevertheless, no direct comparisons can be made as the Percoll gradient used in this assay and those reported for wheat and spinach are not identical in composition, the number of steps, or the centrifugation conditions, which are the key variables that define the behavior of a Percoll gradient. The level of chloroplast contamination obtained here in the phase corresponding to 20% Percoll was comparable with that obtained with ConA affinity columns, but remarkably, the haustoria yield was 5.6-fold higher.

Although a wide range of methodological variations can be applied to Percoll gradients to improve their efficiency (Amersham-Biosciences, 2001, Ruizromero et al., 1995), small variations were introduced to this method in order to maximize the recovery of *Pst* haustoria. Based on an extensive literature review, I decided to shorten the time of centrifugation and increase the concentration of
Percoll on the bottom step to 40%. The increase in the Percoll density at the bottom step of the gradient would impose a density barrier to the haustoria, forcing them to stay in the 20% phase. In turn, these necessitate a larger volume of the 20% phase to prevent saturation. These changes resulted in ~75% of the total haustoria being contained in the 20% Percoll step of the gradient. A striking ~8.6-fold increase in the number of recovered haustoria was achieved using this three-step gradient, although the number of chloroplasts per haustoria was on average 2.5 times that of the ConA-affinity method. Overall, discontinuous Percoll gradients gave much higher yields than the ConA column technique, with obvious advantages in the ease of sample preparation. Depending on the final application, different experimental strategies can be used to enrich haustoria, either minimizing the amount of contaminant chloroplasts, or to maximize haustorial recovery while slightly sacrificing purity.

This part of the study emphasizes that there is no simple way to determine the optimal parameters to apply discontinuous Percoll gradients to haustorial enrichment. Chloroplasts were the major contaminant and haustoria could be only partially separated from them on the basis of density. Apart from the concentration of Percoll in the steps of the gradient, technical aspects such as the time of centrifugation, the volume of the steps of the gradients, and the magnitude of the gravitational force, can strongly influence the distribution of the sample across the gradient. In general, the use of discontinuous gradients gave a rough guide to select the best Percoll concentration for sedimentation of *Pst* haustoria, and to find an equilibrium between maximum recovery of these structures with the least possible contamination. The major disadvantage of utilizing this method is the technical difficulty involved in preparation of the gradients, as it is difficult to get stable sharp-step gradients and therefore reproducible results.

Self-formed continuous gradients offer more flexibility than step gradients and are easier to use. Cells or subcellular particles can be mixed with Percoll to a desired final concentration prior to centrifugation, and will band isopycnically as the gradient is formed *in situ* (Amersham-Biosciences, 2001). Based on the empirical information derived here of haustoria sedimentation on discontinuous gradients, and the density patterns described on continuous Percoll gradients (Amersham-Biosciences, 2001), I tested self-formed continuous gradients of 20% and 30% Percoll. While ~90% of haustoria were retained on the top half of 30% Percoll gradients, haustoria were spread more evenly amongst the fractions on 20% gradients and a large percentage pelleted at the bottom of the tube with cell debris and other contaminants. Chloroplasts were always the main contaminant in both cases, although their distribution patterns were different. On 20% Percoll gradients, chloroplasts
were mostly at the bottom of the gradient, allowing the recovery of haustoria from the top fractions with very low levels of contamination. On 30% Percoll gradients, the chloroplasts were distributed throughout the gradients but showed two bands of high intensity. The upper of these was overlapped to the fractions of maximal haustoria content, and was therefore impossible to avoid while harvesting haustoria. Approximately one third more haustoria were recovered from 30% than 20% Percoll gradients, but they were four times more contaminated from the higher concentration matrix. Thus, the 30% gradients helped to maximize haustoria recovery, while 20% gradients helped to minimize contamination. In both cases, the efficiency was higher than the ConA affinity method. Additionally, these results raise the possibility that the application of consecutive gradients of 30% and 20% Percoll might substantially improve the purity of isolated haustoria while maintaining higher yields than the ConA affinity method. Self-formed continuous Percoll gradients can be modified in various ways by changing the initial concentration of Percoll, the length of the gradient, and centrifugation speed and time. It is possible that variation of one or more of these factors could improve the yield of haustoria and diminish the chloroplast contamination level. However, for the scope of my project, enriched suspensions of haustoria produced by the Percoll gradients described here were used successfully for the generation of transcriptomic data from *Pst* haustoria (Garnica et al., 2013) and were the base for further haustoria purification by flow cytometry (described in Section 4.3.3). Thus, no further methodological variations were evaluated.

Interestingly, additional experiments with 30% Percoll self-formed continuous gradients were used to determine that there are three peaks in the distribution of haustoria, equivalent to different densities. It is interesting to note that the principle of isopycnic centrifugation is the assumption that cells of the same type have the same density, and so will sediment at the same level of the gradient (Patel, 2000). However, this was not case for *Pst* haustoria cells as not all banded at the same level. Studies on yeast cells have shown that they undergo periodic fluctuations in density during the cell division cycle (Mitchison, 1958). This fluctuation has been explained by the fact that dry mass increases continuously throughout the cell cycle, whereas the increase in cell volume has been reported to occur during the latter half of the cycle (Mitchison, 1958, Hartwell, 1970). Therefore, if mass and volume changes do not proceed coordinately, density variations would occur. Also, the fact that the vacuole is more prominent in the cells of lightest density suggests that this organelle might account for the disproportionate mass and volume changes (Hartwell, 1970). In the bacterial model *Escherichia coli* morphological and physiological changes including, changes in the plasma membrane composition, accumulation of storage compounds such as glycogen and polyphosphate and the decrease in polyamines, may lead to an increase in the cell buoyant density (Makinoshima et
al., 2002). Recent studies on the temporal and spatial development of *Pst* haustoria in wheat showed that haustorial size and morphology changed substantially between 0 and 5 dai (Sorensen et al., 2012). Haustoria increase their surface area more than 10-fold from small spherical haustoria to branched mature haustoria. Electron microscopy of the central part of the haustorial body showed that more mature haustoria appear to be rougher and grooved (Sorensen et al., 2012). Ultrastructural studies on barley powdery mildew have shown that the haustorial space occupied by the vacuole varies according to the developmental stage of the haustorium (Hippe, 1985). All such developmental changes in haustoria could promote changes in density and therefore the distribution of these cells on a density gradient. Other factors such as the presence/absence of the extrahaustorial matrix and extrahaustorial membrane on the isolated haustoria could influence their buoyant density. Thus, further studies are necessary to clarify the origin of the variation in haustoria density detected by the continuous gradients of Percoll.

Discontinuous and continuous Percoll gradients were useful here to substantially enrich haustoria from homogenates of infected tissue, providing a good alternative to the ConA-affinity method. Unfortunately, it was not possible to completely eliminate contaminating chloroplasts on any of the gradients, and the contamination levels were always comparable or higher than those obtained by ConA-affinity columns. Nevertheless, in a scenario where wheat mesophyll cells contain approximately 150 chloroplasts per cell (Boffey and Lloyd, 2011, Wardley et al., 1984), they constitute only 50% of the total number of leaf cells (Jellings and Leech, 1982), and not all of them are invaded by a haustorium, the haustoria:chloroplast ratio in raw homogenate of infected tissue will be far greater than 1:150. This ratio was reduced at least by half using Percoll gradients, while recovering the vast majority of haustoria contained in the original sample. It is very important to note that although large efforts were taken to obtain uniform stripe rust infections each time the experiments were repeated, variations of the intensity of the infection were inevitable, and this impacted both yield and purity of haustoria.

4.4.2. Purification of haustoria by fluorescence-activated cell sorting (FACS) and evaluation of their viability post sorting

In wheat rusts, spores collected from erupting pustules on infected leaves are the only fungal tissues that can be obtained virtually free of contaminating host material. Infectious hyphae, haustorial mother cells and haustoria are embedded in the host tissue and the complete purification of these structures has not been possible. The haustorium probably represents the single most important
structure for understanding the host-pathogen interaction (Voegele and Mendgen, 2011, Catanzariti et al., 2007) but so far specific transcriptomic and proteomic studies have been hampered by the inability to separate it from contaminating host material. Techniques such as the ConA-affinity method have helped to deepen the study of haustoria of different rust pathogens (Hahn and Mendgen, 1992), however its utility is limited due to the low yield and purity of the haustorial preparations. The remarkable increase in the number of sequenced fungal genomes and transcriptomes has facilitated the molecular studies of haustorial tissue regardless of the level of contamination that accompanies the isolated material (Wessling et al., 2012, Fernandez et al., 2012, Xu et al., 2011, Vieira et al., 2012). However, these resources do not yet exist for many important pathogens, as was the case for *Pst* at the beginning of this project.

A major application of fluorescence-activated cell sorting is to separate cells according to subtype, epitope expression or fluorescent markers from mixed cell suspensions or homogenates of complex tissues. This technique has been used extensively in biomedical research, clinical microbiology, environmental microbiology (Porter et al., 1997, Davey and Kell, 1996, Alvarez-Barrientos et al., 2000, Brown and Wittwer, 2000), and to a lesser extent in plant research (Yanpaisan et al., 1999, Ochatt, 2008). Flow cytometry has potential applications in plant phytopathology including the characterization of genome sizes of fungal and oomycete populations (Vercauteren et al., 2011, Eilam et al., 1994), multiplexed pathogen detection (Iannelli et al., 1996, Iannelli et al., 1997, Chitarra et al., 2002) and the monitoring of the viability, culturability and gene expression of plant pathogens, amongst others (Chitarra et al., 2006, D'Hondt et al., 2011). Takara et al (2009) used FACS for purifying primary hyphae of *Colletorichum higginsianum* from infected tissue. In this part of the study I showed that FACS can be applied to purification of *Pst* haustoria, taking advantage of the affinity of ConA to rust fungi, resulting in much higher levels of purity than previous methods.

Cell sorting was based on the selective staining of haustoria by ConA fused to the fluorophore Alexa 488. This fluorophore and the chlorophyll contained in chloroplasts can be excited with the same laser but emit different wavelengths, making haustorial and chloroplast populations clearly distinguishable and separable by FACS. The high specificity of ConA for haustoria but not for other fungal tissues or plant cell debris makes this technique highly selective. Haustoria samples isolated by FACS exhibited a purity close to 98% and so there was a dramatic change in the haustoria:chloroplast ratio, from ~1:60 before sorting to ~49:1 afterwards. The maximum level of purity obtained by the ConA-affinity method and Percoll gradients (continuous 20% Percoll) were ~6.6% and ~5.8% respectively, not nearly comparable with the results obtained by FACS. The time
required for sorting a given number of target cells is directly related to their concentration in the input sample (Fisher et al., 1998). Thus, although FACS could be directly applied to raw homogenates of infected tissue, the previous enrichment of haustoria by self-formed continuous Percoll 30% gradients conveniently minimizes the sorting time to approximately two hours (the entire isolation process takes ~5.5 h). There is no literature available on the effect of releasing haustoria from the host cells on the haustoria survival rate. Presumably, very long periods of manipulation after release could greatly affect their viability or indeed many other aspects of metabolism or gene expression. As will be discussed later, the time required to isolate haustoria using Percoll gradients and FACS does not appear to have a significant impact on haustoria viability.

On average, only 25% of the total haustoria contained in the pre-sorted fraction were recovered after FACS. This loss can be attributed to different causes. Apart from very high green fluorescence (selecting for Alexa 488) and non-red fluorescence (to discard chloroplasts), additional parameters such as forward and side scatter of light were taken into account during the sorting process. Preliminary assays combining sorting and microscopy were necessary to optimize the values of all parameters to avoid contamination of the sorted haustoria population. In this way, the sorting parameters were set very stringently, and most likely this contributes to the loss of haustoria during sorting. It has been proposed that the affinity of ConA for isolated haustoria comes from the proliferation of ConA-binding carbohydrates present in the extrahaustorial matrix (Hahn and Mendgen, 1992). This matrix presumably becomes exposed during the release of haustoria from infected tissue as haustoria lose their extrahaustorial membrane (Hahn and Mendgen, 1992). The extent to which haustoria lose their extrahaustorial membrane and extrahaustorial matrix during the homogenisation method applied in this study is unknown. One possibility is that enriched suspensions of haustoria from Percoll gradients are a mixture of haustoria with none, partial or full presence of the extrahaustorial membrane and matrix. Thus, this phenomenon would also have a direct effect on the degree of fluorescent labeling and therefore on the haustoria yield after FACS. Nevertheless, the yield obtained in these assays was sufficient for subsequent analyses as will be discussed in Chapter 5.

Some factors during the isolation process could potentially affect the physiological status of haustoria. For example, once haustoria are released from infected tissue it can take up to 5.5 h until the sorting process is completed. Also, the sorting process in itself involves a pressurization step and a very brief laser exposure, factors with unknown effect on haustoria. Hence, the determination of haustoria viability was essential to evaluate their physiological state post-sorting, and thus the utility
of this technique for downstream applications. Classification of microbial cells as viable or non-viable was based for many years on the basis of their ability to reproduce (Postgate, 1969). However, haustoria do not multiply either in vivo or in vitro (Jones, 1974). There are no established methods to assess haustoria viability and only a few reports where the viability of isolated haustoria has been tested (Dekhuijzen, 1966). Current methods of detecting cell viability in a wide range of cells exploit the properties of healthy cell membranes to exclude dye, or alternatively, exploit a healthy cell’s metabolic activity to convert a dye to a visible state (Haugland, 2001). Here I applied two methods to assess haustoria viability on pre- and post-sorted samples. The first was a membrane leakage assay using 7-Aminoactinomycin D, which is a DNA intercalating dye expected to stain only those cells with compromised membranes. Microscopic analyses of pre-sorted samples tested with 7-AAD suggested that haustoria remained viable at that stage of the isolation process as they were impermeable to the dye. Conversely, detergent-treated haustoria (the negative control) became permeable to the dye because apparently their membranes were damaged. The application of the test on post-sorted samples indicated that the sorting process did not affect the viability of haustoria, as the dye did not penetrate even after longer incubation periods than are commonly used with this dye. One of the advantages of the use of 7-AAD to test viability is that the fluorescence is stable for up to two hours post fixation because of the high DNA-binding constant and slow dissociation rate (Schmid et al., 1992, Philpott et al., 1996). This reduces the rate of leakage of 7AAD out of the positive cells and makes this test suitable for dead cell discrimination by FACS (Philpott et al., 1996). Thus, it was possible to apply the 7-AAD test on sorted cells and quantify haustoria viability by re-sorting the samples. Remarkably, only 5% of post-sorted haustoria were discriminated as non-viable cells by 7-AAD.

The scarcity of useful vital indicators made it difficult to derive strong conclusions on the viability of sorted haustoria from the 7-ADD test. Also, there is a possibility that a haustorium’s viability may have been compromised even though its plasma membrane integrity is (at least transiently) maintained. Conversely, the integrity of the haustorium membrane may be abnormal, yet the cell may be able to repair itself and still be viable, as shown in yeast cells (Davey and Hexley, 2011). Thus, to complement the results obtained from the 7-AAD tests, a second viability test was applied on pre- and post-sorted haustoria to investigate their capacity to metabolize reagents. CellTracker Orange CMRA is a fluorescent probe that passes freely through cell membranes and measures enzymatic activity, which is required to activate their fluorescence, and cell-membrane integrity, which is required for intracellular retention of the fluorescent product. CMRA contains a chloromethyl group which reacts with glutathione (specifically with its thiol group) in viable cells via glutathione-S-
transferase. The dye remains non-fluorescent until it is cleaved by cytosolic esterases. The cleavage of the acetate and the release of the cell-impermeant fluorescent dye-thioether allows visualization of cytoplasm which is usually relatively uniformly stained. In most cells, glutathione levels are high (up to 10 mM) as it has an important role as an antioxidant, eliminating free radicals and toxic oxidants from the cell. Therefore, in most eukaryotic cells glutathione is a vital molecule and its presence is an indicator of cell viability (Arafeh, 2008). It was critical to optimize the conditions of this test for haustoria; the establishment of reliable negative controls and microscope configuration were important steps that had to be adjusted before applying it to sorted samples. Special attention was paid to the negative controls as ideally the treatment to kill haustoria should induce loss of metabolic activity while preserving haustorial structures. Mild heating and formaldehyde treatments were used as haustoria killing agents as these treatments typically abolish metabolic activity in eukaryotic cells (Guinan et al., 2010, Thavarajah et al., 2012). Preliminary assays indicated that haustoria were permeable to CMRA as they fluoresced internally, although this was distributed unevenly in the cytoplasm. Interestingly, haustoria present in samples treated with heat or formaldehyde exhibited background fluorescence after the CMRA test was applied. Nevertheless, haustoria from treated and untreated samples were easily distinguishable based on their levels of fluorescence, suggesting that CMRA could discriminate between viable and non-viable haustoria, and could therefore be useful to assess sorted samples.

In the last experiment to assess the viability of post-sorted haustoria by CMRA, I compared freshly sorted and 5 days post-sorted haustoria samples. I hypothesized that haustoria released from their natural environment (infected host tissue), maintained for several hours at ~21°C (the optimal temperature for *Pst* is 17°C), and suspended in a media with no nutritional value (1x PBS) would rapidly lose their metabolic activity and could be useful as an additional negative control. Clear differences in fluorescence between freshly sorted and 5 days post-sorted haustoria were observed. Freshly sorted haustoria remained metabolically active after the isolation process, suggesting these purified cells are viable and amenable to further experimentation. By contrast, the metabolic activity of 5 days post-sorted haustoria declined to levels that could not be detected by the CMRA test, indicating that they had lost viability. Although the available time did not allow me to do proper comparative quantitative analyses to find the exact percentage of viable haustoria after sorting, close to 100% of the haustoria in freshly sorted samples gave positive results in this test. The results of the test agrees with the results obtained with 7-ADD, strongly suggesting that a very small fraction of haustoria result were damaged during the two-phase method for enrichment and purification described in this chapter.
Overall, by using Percoll gradients it was possible to increase dramatically the yield of *Pst* haustoria per gram of infected tissue, facilitating the enrichment of haustoria for further experimentation. So far, sole use of Percoll gradients in the absence of FACS was useful for the generation of haustorial transcriptomic data (Garnica et al., 2013). This method has been successfully adapted to other fungal haustoria-forming pathogens for the generation of transcriptomic data (N. Upadhyaya, CSIRO Plant Industry, personal communication), facilitating the work with non-culturable organisms. FACS applied to enriched suspensions of *Pst* haustoria allowed the purification of viable haustoria which have been successfully subjected to proteomics analyses (Garnica et al, in preparation). Together these two techniques comprise powerful tools to advance the understanding of biotrophic haustoria, the key structure in some of the most devastating pathogens of crop plants.
Chapter 5

Proteomics of *Pst* spores and haustoria as a tool to improve the investigation of effectors

5.1. INTRODUCTION

Fungal proteomics research has advanced dramatically over the past 5 years. The development of powerful technologies such as tandem liquid chromatography-mass spectrometry (LC/MS) (Braaksma et al., 2010, Paper et al., 2007), and the introduction of high-throughput DNA sequencing technologies have revolutionised the collection and interpretation of mass spectra. Although proteomics is defined as the large-scale study of the presence, structure and function of proteins in complex biological samples, additional information such as localization, posttranslational modifications and even quantification of protein levels can be extracted from proteomic data (Gonzalez-Fernandez and Jorrin-Novo, 2012). These possibilities make this technology an integral component of the “omics” systems (Gonzalez-Fernandez et al., 2010, Doyle, 2011). Proteomics approaches can provide clues on where, when, and for what reason signatures of thousands of individual protein species are obtained. Nowadays, high-throughput proteomics is a valuable tool to support genome annotation, particularly the in silico gene models predicted from the increasing
number of fungal genome sequencing projects. The collection of data from different sources (tissues/structures) is an efficient strategy to increase the coverage of the ‘peptide-validated’ genome, while providing information on the proteome heterogeneity between developmental stages (Wright et al., 2009).

Proteomics studies involve analysis of complex protein mixtures such as tissue homogenates or cellular lysates. These biological samples contain large numbers of proteins with a range of physicochemical properties present in widely differing concentrations (Dauly et al., 2006). This poses a number of limitations on the sensitivity, resolution, coverage and quantification of proteomics techniques, strongly limiting the possibility of uncovering the entire complement of proteins, and making this discipline much more challenging than other nucleic acid-based “omics” technologies. The sampling and extraction protocol must be optimized for the biological system and the objectives of the research (Jorrin-Novo et al., 2009, Gonzalez-Fernandez et al., 2010). The workflow usually includes the following steps: experimental design, sampling, sample preparation, protein extraction, protein separation, peptide preparation, mass spectrometry (MS), identification of peptides and derivative proteins, and statistical analysis of the data (Gonzalez-Fernandez and Jorrin-Novo, 2012). The fractionation of the protein mixture prior to MS decreases the sample complexity, allowing more accurate peptide identification and detection of lower abundance proteins (Krause et al., 1999). Sodium-dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) in a 1- or 2-dimensional format (1-DE, 2-DE) is the most widely used separation technique prior to MS analysis (Petriz et al., 2012). Although 1-DE and 2-DE offer good protein separation, these techniques are labour intensive and it is difficult to automate the fractionation, excision and digestion processes. Other drawbacks include poor reproducibility, the inability to detect low abundance and hydrophobic proteins, poor sensitivity in identifying proteins with very high or very low pl values, and difficulty in detecting proteins with very low (Mr < 10 kDa) or very high (Mr > 150 kDa) molecular masses (Chandramouli and Qian, 2009, Monteoliva and Albar, 2004). These limitations of gel-based fractionation methods have motivated the development of gel-free techniques based on separation of proteins or peptides by liquid chromatography (LC), including two-dimensional LC based on a high-performance chromatofocusing in the first dimension followed by high-resolution reversed-phase chromatography in the second (Pirondini et al., 2006, Gonzalez-Fernandez et al., 2010, Neverova and Van Eyk, 2005). These techniques improve quantitative and qualitative proteomic analyses and expand the possibilities of these strategies. For example, the OFFGEL fractionator (off-gel electrophoresis device commercialized by Agilent Technologies), replaces one dimension of 2-DE since it separates peptides and proteins according to their isoelectric points on
immobilized pH gradient strips in liquid phase with very high resolving power (Michel et al., 2003, Heller et al., 2005, Horth et al., 2006, Ros et al., 2002). The combination of this method with LC-MS provides a good alternative for proteomics analyses (Geiser et al., 2011).

Irrespective of the choice of a given proteomic separation technique, a mass spectrometer is always required for peptide identification. It allows proteome profiling from a qualitative (and sometimes quantitative) point of view but more importantly, it allows the identification of protein species (Chandramouli and Qian, 2009). Mass spectrometers consist of an ion source, the mass analyzer, and an ion detection system. Analysis of proteins by MS occurs in three major steps: (a) protein ionization and generation of gas-phase ions, (b) separation of ions according to their mass to charge ratio, and (c) detection of ions (Mann et al., 2001, Chandramouli and Qian, 2009). A mass spectrometer’s resolution, mass accuracy and sensitivity are the primary considerations for determining the performance of the machine and therefore, the quality of data generated from the analysis of protein samples (Chernushevich et al., 2001). The peptide masses resulting from the MS analysis are matched with the peptide fingerprints of known proteins or theoretical ones generated from gene sequences, using search engines (e.g., Mascot, Sequest, Comet, X!tandem) (Chandramouli and Qian, 2009, Gonzalez-Fernandez et al., 2010). Results are scored according to a scheme specific to each search engine that also depends on the database used for the search. Thus, the different search engines do not give identical results as they are based on different algorithms and scoring functions, making comparison and integration of results from different studies or experiments very challenging (Carr et al., 2004). Additionally, the peptide identification process is computationally intensive and time-demanding. The association of identified peptides with their precursor proteins is a critical and very difficult step since many peptides are common to several proteins, thus leading to ambiguous protein assignments. For this reason it is essential to have a statistical tool that is able to assess the validity of the protein inference and associate a probability to it (Chandramouli and Qian, 2009). The number of proteins identified from MS data is also dependent on the availability of genomic DNA or EST sequences and is often the major bottleneck in proteomics research on unsequenced “orphan” organisms, as is the case for most plant fungal pathogens (Gonzalez-Fernandez et al., 2010). This situation is expected to change rapidly since the advent of rapid DNA sequencing methods is greatly accelerating the availability of genomes of plant pathogenic fungi as well as those of their host plants, which enormously benefits proteomic studies in these pathosystems.
The first proteomics studies on fungal pathogens were carried out in order to understand their interactions with plants by searching for resistance-related proteins. The identification of the avirulence gene product Avr9 in the tomato-Cladosporium fulvum pathosystem by the purification of apoplastic fluid from susceptible plants was the pioneer work in this field in 1988 (Schottenstoma and Dewit, 1988). Since then, numerous studies have helped our understanding of plant–fungal interactions, fungal pathogenicity, and virulence (Rep et al., 2004, Bhadauria et al., 2010a, Gonzalez-Fernandez and Jorrin-Novo, 2012, Gonzalez-Fernandez et al., 2010). In this way, the first effector (Six1) from the root invading pathogen Fusarium oxysporum was identified in the tomato pathosystem (Rep et al., 2004), and some other protein effectors have been identified from different fungi (Rep, 2005). Currently, the availability of fully sequenced fungal phytopathogen genomes has enabled the prediction of large numbers of effector gene candidates. However, the validation of these predictions using proteomics has mostly been achieved in culturable fungi (see reviews (Gonzalez-Fernandez and Jorrin-Novo, 2012, Gonzalez-Fernandez et al., 2010)). In the case of haustoria-forming biotrophic fungi, the identification of secreted effectors through proteomics is mostly limited to studies in planta or haustoria-enriched samples. In planta studies are much more complicated due to the presence of both proteomes (plant and pathogen), whereas the pathogen’s secreted proteins are a tiny proportion of the total sample and hence hardly detectable (Bindschedler et al., 2009, Song et al., 2011, Gonzalez-Fernandez et al., 2010).

Proteomic studies on conidia, the major spore type of Ascomycete fungi, have helped to understand spore germination and development in the initial stages of the infectious lifecycle (Sulc et al., 2009, Noir et al., 2008, Bhadauria et al., 2010b). For example, proteins from Blumeria graminis f. sp. hordei conidia were resolved by 2-DE and peptides identified by MS. A total of 123 distinct fungal gene products were identified, most with predicted functions in carbohydrate, lipid, or protein metabolism (Noir et al., 2008) and were correlated with previously reported transcriptomic profiles of the same developmental stage (Both et al., 2005a, Both et al., 2005b). In the rice pathogen Magnaporthe oryzae, the conidial proteomes of the wild type and the com1 (a putative transcriptional regulator of conidial morphology) mutant strain were compared (Bhadauria et al., 2010b). 2-DE combined with matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectroscopy analysis allowed the identification of 31 proteins with altered levels, the genes for which are potentially regulated by the Com1 protein. This study suggested that the Com1p may play roles in transcriptional reprogramming of genes implicated in melanin biosynthesis, carbon and energy metabolism, structural organization of the cell, lipid metabolism, and amino acid metabolism. A comparison between germinating and dormant uredospores of the bean pathogen Uromyces
appendiculatus revealed that after germination, there were significant changes in the relative concentration of proteins involved in energy production and nucleus reorganization (Cooper et al., 2007). These changes suggest that uredospores have high energy demands and require structural proteins during germination, consistent with a metabolic transition from dormancy to germination. These examples illustrate the utility of proteomics for identifying metabolic changes in the early stages of infection. However, proteomic studies on conidia are still scarce for biotrophic plant pathogens.

So far, no proteomics studies have been reported for Pst. Thus, in an effort to investigate the molecular mechanisms underlying the pathogenic structures of Pst more closely, proteomics analyses of Pst spores and haustoria were undertaken. In this chapter, I describe recently obtained data from Pst infectious and biotrophic stages.

5.2. MATERIALS AND METHODS

5.2.1. Sample processing and protein extraction from ungerminated and germinated Pst uredospores.

To prepare material from germinated spores, 300 mg of freshly collected uredospores were tapped onto 300 ml of MilliQ water contained in a glass petri dish (~15 cm diameter) and were incubated for 15 h at 9°C in darkness. Once germination was verified by bright field microscopy, the water was removed by filtration through a 15 µm nylon mesh, and the remaining moisture was removed as much as possible with paper towel. Dried samples were ground in liquid nitrogen and stored at -80°C until processed for protein extraction. To prepare material from ungerminated spores, 300 mg of dried spores were ground extensively to a very fine powder in liquid nitrogen and stored at -80°C until processed for protein extraction.

Proteins were extracted from germinated and ungerminated spores following the TCA/acetone extraction method described in (Vincent et al., 2012b). Briefly, frozen powdered spore material was resuspended in 3 ml of chilled extraction buffer (10% (w/v) trichloroacetic acid (TCA), 0.07% β-mercaptoethanol in acetone) and incubated overnight at -20°C. The sample was centrifuged at 12,000 g at -10°C for 30 min to pellet proteins. The pellet was resuspended in 4.5 ml of chilled wash buffer (0.07% β-mercaptoethanol in acetone) and incubated for 2 h at -20°C with regular manual agitation. The washing step was repeated two more times. In the last wash the supernatant was
removed and the tubes were immediately closed, snap frozen in liquid nitrogen and stored at -80°C until they were processed for protein analysis. The subsequent protein analyses were done in collaboration with Dr Thomas Stoll and Dr Jeffrey Gorman at the Protein Discovery Centre, Queensland Institute of Medical Research.

5.2.2. OFFGEL fractionation of spore proteins.

Germinated and ungerminated spore proteins were separated from insoluble spore material by solubilising proteins in UTC buffer (7M urea, 2 M thiourea, 4% (w/v) CHAPS, 1% dithiothreitol (DTT)) at 4°C for 1 h using a rotary shaker. Proteins solubilized in UTC were verified and quantified on a SDS-PAGE gel. For isoelectric point-based protein separation, the 3100 OFFGEL Fractionator (Agilent Technologies, Böblingen, Germany) was used with a 24-well set-up according to the manufacturer’s protocol. The sample was diluted with OFFGEL peptide sample solution and loaded in each well for electrofocusing. Twenty four peptide fractions were sampled before reduction with 45 mM DTT and alkylation with 100 mM iodoacetamide, before digestion with trypsin to produce peptides for MS.

5.2.3. Protein extraction from FACS-purified Pst haustoria

Proteins from FACS-purified haustoria were processed by two methods. In method A, haustorial proteins were extracted by boiling ~1.5 x 10^6 haustoria cells in SDS extraction buffer (50 mM Tris-HCl pH 8.0, 0.1 M EDTA, 2% SDS, 10 mM DTT) for 5 min at 95°C. Ten volumes of chilled 10% trichloroacetic acid in acetone (~20°C) were added to the protein extract and the mixture was incubated overnight at ~20°C. The sample was centrifuged at 15,000 g for 10 min at -10°C. The supernatant was removed and the pellet was washed in the same volume of chilled acetone and incubated at -20°C for at least 10 min. After centrifuging the sample at 15,000 g for 5 min, as much supernatant was removed as much as possible, and the tube was immediately closed to avoid desiccation. The proteins were dissolved in 6 M urea, reduced, alkylated and double digested with LysC and trypsin. In method B, ~1.5 x 10^6 haustoria cells were resuspended in 6 M urea and incubated overnight at 4°C to lyse cells and solubilise proteins. The proteins were reduced, alkylated and double digested with LysC and trypsin.

5.2.4. Capillary liquid chromatography and mass spectrometry of Pst peptides
Peptides from each sample (or individual fractions from fractioned samples) were separated by reversed-phase chromatography using an octadecyl carbon chain (C18) column. Chromatograms were recorded at the wavelength of 214 nm. Peptide fractions were collected and the eluate was electrosprayed into an LTQ Orbitrap Velos high-resolution mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) through a Proxeon nanoelectrospray ion source. All of these processes were performed at the Protein Discovery Centre, Queensland Institute of Medical Research.

5.2.5. Mass spectrometry data analysis

For protein identification, the Pst reference transcriptome obtained from de novo assembly of combined 454 data from haustoria and germinated spores (Chapter 3 and Data S4 in Garnica et al., 2013) and the set of predicted Pst effectors (Chapter 2 and Data S3 in Garnica et al., 2013) were combined with the reversed sequences (decoy database) and sequences of widespread contaminants, such as human keratins, trypsin and bovine serum albumin. Using the Mascot search engine, carbamidomethylation (C) was set as a fixed modification while oxidation (M) and deamidation (N, Q) were set as variable modifications. Initial peptide mass tolerance was set to 20 ppm and fragment mass tolerance was set to 0.8 Da. Two missed cleavages were allowed and the minimal length required for a peptide was eight amino acids. Monoisotopic experimental mass values, peptide charges 2+ and 3+, and no error tolerance were additional configured parameters in Mascot. Data from the same tissue were merged. Peptide and protein probabilities were assigned with Scaffold using the built-in prophet tools from the Trans-Proteomic-Pipeline to accept peptide and protein identifications (Searle, 2010). The settings used for high-confidence protein identifications were detection of two unique peptides from a single protein, protein probability 99%, and peptide probability 95%.

5.2.6. Assessment of effector candidate gene expression in ungerminated spores by reverse transcriptase-PCR (RT-PCR)

Total RNA was isolated from enriched suspensions of haustoria purified on 30% Percoll discontinuous gradients, and from ungerminated spores. The tissues were ground in liquid nitrogen and extracted using the QIAGEN (Doncaster, Australia) Plant RNeasy kit according to the manufacturer’s instructions. For cDNA preparation, 2 µg of total RNA was mixed with 1 µl of oligo(dT)18 and MilliQ water up to 11 µl and heated at 70°C for 10 min, followed by cooling on ice for one minute. DTT was added to the RNA to a final concentration of 1 mM together with 1 µl of dNTPs
(each at 10 mM), 4 µl of 5X Superscript II buffer (supplied with Superscript kit, Invitrogen) and 0.5 µl of SuperScript II reverse transcriptase (200 units/µl). After incubation for 1 h at 42°C, the reaction was stopped by heating at 70°C for 15 min and 1 µl of RNase H (1U/µl) was added to each tube and incubated for 20 min at 37°C. Target cDNAs were PCR amplified using a dilution of 1:15 (1:5 for haustoria samples) of the cDNA as template, specific forward and reverse primers (a list of primers is in Supplementary Table 2.1), and 2X PCR Master Mix (Promega). PCRs consisted of 30 cycles of denaturation at 95°C, annealing temperature of 58°C and extension at 72°C for 45 sec.

5.3. RESULTS

5.3.1. Extraction of proteins from Pst spores and FACS-purified haustoria

Proteins from both ungerminated and germinated spores were extracted using the TCA/acetone method. In both cases the disruption of the tissues was critical to extract as much protein as possible. Germinated spores were easily ground in liquid nitrogen to a fine powder as the germination process makes them easier to break. On the other hand, ungerminated spores which are more difficult to disrupt were processed using two methods. In method A, ungerminated spores were extensively ground in liquid nitrogen to a very fine powder and were processed as for germinated spores. In method B, ungerminated spores were mixed with sterile acid-washed sand (1:3 w/w), then ground in liquid nitrogen to a very fine powder, and resuspended in Tris-HCl buffer. After centrifugation, the supernatant was recovered and five volumes of TCA in acetone were added to precipitate proteins, and the pellet was washed as in the TCA/acetone extraction method. Total protein concentrations were determined by Bradford assay (Bio-Rad) and the protein profile of these extracts were verified by SDS-PAGE electrophoresis as shown in Figure 5.1.
Figure 5.1. SDS-PAGE of protein extracts from germinated and ungerminated spores. 10 µg of protein extract from germinated and ungerminated spores were verified by SDS-PAGE (Coomassie stained). (1) germinated spores, (2) ungerminated spores proteins extracted by method B, (3) ungerminated spores proteins extracted by method A, (M) 5 µL of Biorad Kaleidoscope Precision Plus Protein Standards marker.

Both methods for extracting proteins from ungerminated spores appeared to work similarly as the protein profiles did not show major differences (Figure 5.1, lanes 2 and 3), and the total amount of recovered protein was fairly similar (~2.5 mg). Thus, method A was chosen for subsequent analysis as this was the same method used for germinated material, and so would be advantageous for comparative studies. It was evident from the gel profiles that both germinated and ungerminated proteins extracts were complex mixtures of proteins with variable concentrations. To reduce the complexity of these samples, 1 mg of protein from each sample was OFFGEL fractionated before LC-MS analyses. Proteins were electrofocused according to their isoelectric points along a pH gradient (3-10 range) and separated into 24 fractions. Each fraction was recovered and treated with DTT to reduce disulfide bonds. Proteins were then alkylated on cysteines to prevent formation of disulfide bonds, and digested with trypsin into peptides before LC-MS.

Mechanical disruption methods were not applied to haustorial samples because of the low number of FACS-purified haustoria cells. Instead, haustoria were chemically disrupted with SDS buffer and proteins were precipitated with TCA/acetone. A second approach was also tested, disrupting haustoria cells and solubilizing proteins in a single step by the use of 6 M urea. Less than 1 µg of total protein was recovered from each extraction method, so these samples were not fractionated by OFFGEL before reduction, alkylation and digestion.
5.3.2. *Pst* transcriptomic resources used for protein identification

The transcriptomic resources developed previously in this research, including the combined *de novo* transcriptome from haustoria and spores 454 data (12,282 transcripts, Chapter 3), as well as the 437 predicted effector candidate genes (Chapter 2 and (Garnica et al., 2013)) were used as the search databases. The open reading frames of the 437 predicted effector candidates were used as the primary database (effector database, EDB). The transcripts from the combined transcriptomes were translated into all six frames, and all sequences larger than 6 amino acids were included. A total of 736,839 protein sequences (including 8 contaminants: keratin, trypsin, BSA) were used as the secondary reference database (transcriptome database, TDB).

5.3.4. Peptides/proteins identification through Mascot software

The search databases were uploaded to the Mascot server of the Walter and Eliza Hall Institute of Medical Research (Melbourne, Australia). Mascot identifies proteins by interpreting mass spectrometry data (Perkins et al., 1999). The software takes the molecular weight of peptides and compares them against a database of known peptides. The program cleaves every protein virtually in the database according to the recognition site of the protease used for digestion, and calculates the theoretical mass for each peptide. Mascot then computes a score based on the probability that the peptides from a sample match those in the selected protein database. The more peptides Mascot identifies from a particular protein, the higher the Mascot score for that protein. The mass spectra generated from the three *Pst* tissues were matched to possible protein identities by comparing them to EDB and TDB databases. Using output data from Mascot, Scaffold software was used to generate peptide identification and protein identification probabilities. In all cases, a minimum of two significant peptides, >95% peptide probability, and 99% or better chance of a protein being correctly identified, were the parameters used to filter data. Results are summarized in Table 5.1.
Table 5.1. Proteins identified from *Pst* proteomic data generated from germinated spores, ungerminated spores and FACS-purified haustoria.

<table>
<thead>
<tr>
<th><em>Pst</em> tissue</th>
<th>Sample treatment</th>
<th>Unique peptides</th>
<th>Proteins identified in TDB</th>
<th>Total unique peptides</th>
<th>Total proteins identified in EDB</th>
<th>Total proteins identified in TDB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Germinated spores</strong></td>
<td>No fractioned sample</td>
<td>7,897</td>
<td>1,501</td>
<td>17,010</td>
<td>71</td>
<td>2,312</td>
</tr>
<tr>
<td></td>
<td>Fractioned sample</td>
<td>14,545</td>
<td>1,952</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Unergminated spores</strong></td>
<td>Fractioned sample</td>
<td>21,014</td>
<td>2,791</td>
<td>21,014</td>
<td>122</td>
<td>2,791</td>
</tr>
<tr>
<td><strong>Haustoria</strong></td>
<td>Disruption method A</td>
<td>404</td>
<td>84</td>
<td>1,009</td>
<td>25</td>
<td>190</td>
</tr>
<tr>
<td></td>
<td>Disruption method B</td>
<td>795</td>
<td>151</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The first pathogenic stage subjected to proteomics was germinated spores. Since the complexity of the sample was unknown, LC-MS was applied to both OFFGEL fractionated and non-fractionated samples. The number of peptides identified was greater when the sample was first subjected to OFFGEL chromatography. Given that the protein profiles of ungerminated spores were comparable in complexity to germinated spores (Figure 5.1), only prefractioned samples were subjected to LC-MS. On the other hand, haustoria proteins from the chemical disruption methods were analyzed by LC-MS without prior fractionation. Both methods generated a number of peptides and allowed the identification of proteins, although method B appeared to be slightly more efficient since it allowed the identification of almost twice as many proteins as with method A (Table 5.1).

A total of 3,080 proteins and 150 effector candidates were detected after matching the mass spectra from all samples against the TDB and EDB respectively (Note: due to the very large size of this dataset, it was not included as an appendix in thesis. However, all of the data are available at the data repository of Rathjen laboratory and can be made available upon request. Additionally, the original proteomic data are being submitted to a public proteomics data repository and will available to the public upon the publication of this study). Some proteins were found in more than one pathogenic stage, as shown in Figure 5.2.
Figure 5.2. Proteins in Pst haustoria and spores. Mass spectra data were matched against (A) the EDB and (B) the TDB. Some proteins were detected in more than one pathogenic stage as illustrated by the Venn diagrams.

The amount of haustorial tissue available for proteomics was only a fraction of that of spores. This impacted directly on the number of peptides generated and hence the proteins that could be identified. Nevertheless, at least 25 candidate effectors and approximately 190 proteins were detected from haustoria. Due to time constraints the identity and functional categories of the detected proteins in all tissues could not be included in this thesis. However, further analyses of the data using the recently assembled genome of Pst 104 E137 A- and the derived set of gene models (W. Jackson, The Australian National University, personal communication) are currently in progress.

5.3.5. Candidate effector proteins identified by proteomic analyses

Almost 34% (n = 437) of all candidate effector proteins were matched by peptide spectra from these experiments. Of these, 42 effector proteins were at least partially annotated with respect to function, as described in Chapter 2 (Table S2 in Garnica et al., 2013). These included the predicted functions or functional domains of glycoside hydrolases (identifed 13 of the 15 predicted in the whole set), detoxifying enzymes (9 of 10) and peptidases (2 of 3), all of which were detected in spore samples, particularly in ungerminated spores. The genes for all of the 25 effector proteins in detected haustoria were strongly expressed in this pathogenic stage, and 19 of them showed statistically significant up-regulation compared to germinated spores (Table S2 in Garnica et al., 2013). However, the genes for 19 of the 71 effectors detected in germinated spores were not expressed in this pathogenic stage, but interestingly, were significantly up-regulated in haustoria. Furthermore, 25 candidate effector proteins identified in ungerminated spores were not predicted to be expressed in germinated spores. As germinated spores are a metabolically active derivative of
ungerminated spores, it is possible that these genes were also not expressed in the dormant spores. To explore this idea further, the genes for seven candidate effector proteins detected in ungerminated spores were amplified by reverse transcriptase-PCR (RT-PCR) from ungerminated spores and haustoria cDNA (Figure 5.3). Four of these were predicted from transcriptomic data to be non-expressed in germinated spores, and three were highly expressed in this pathogenic stage.

### Figure 5.3. Effector gene expression in ungerminated Pst spores and haustoria.

Seven genes encoding candidate effector proteins detected in ungerminated spores were chosen and their expression patterns were analyzed by RT-PCR. cDNA synthesized from the mRNA extracted from ungerminated spores (SP) and haustoria (H) was used as template. Tubulin protein was also found in ungerminated spores, and expression of this gene was used as a control for amplification. From the expression data, the genes could be classified into two groups: 1. Those with both the protein and mRNA present in a given pathogenic stage, and 2. Those with the protein present, but the mRNA absent.

Messenger RNAs from candidate effectors genes that were not detected by RNA-seq in germinated spores were also not detected by RT-PCR in ungerminated spores. Conversely, mRNAs from candidate effector genes detected by RNA-seq in germinated spores were also detected in ungerminated spores, as shown in Figure 5.3. These findings suggest that effector proteins and/or their messenger RNAs are synthesized distally and transported for deposition in spores during the sporogenesis process, but the where they are synthesized and how they are transported to spores instead of being secreted straight after synthesis requires further investigation.

## 5.4. DISCUSSION
High-throughput LC-MS has revolutionized the proteome analysis of different organisms, including obligate pathogens and parasites that pose the additional challenge of being uncultivable (Florens et al., 2002, Hall et al., 2005, Smolka et al., 2003, Wang et al., 2007, Nilsson et al., 2010, Cooper et al., 2007, Nirmalan et al., 2004). Deciphering the proteomes of these organisms has given valuable insights into their obligate nature, and offered a unique means of determining not only protein abundance, but also subcellular localization and post-translational modifications. This last part of my PhD project aimed to decipher the proteomes of key pathogenic life stages of *Pst* to validate the predictions of my previous transcriptomic analyses (Garnica et al., 2013), and to gain insight into the possible roles of the effectors during disease. Ungerminated and germinated spores are key life stages of *Pst*, firstly as they make the contact with the host plant, and secondly upon successful germination leading to elaboration of metabolic programs to penetrate the leaf. On the other hand, haustoria seem to be responsible for the nutrient uptake and delivery of effectors, two critical roles in pathogenicity. Here I studied all three of these structures using a proteomics approach. Due to the time constraints of the project, only generation of the data and a very preliminary analysis of its contents could be included in this thesis. Nevertheless, more than 3000 proteins predicted from the assembled transcriptome (Chapter 3) and a significant proportion of effector candidates were detected in the proteomics datasets.

Differences in the structure and abundance of each pathogenic stage necessitated different strategies for processing prior to LC-MS analysis. Special attention was paid to cell disruption and protein extraction procedures, as proteins can be detected only if they are first extracted and solubilized (Gonzalez-Fernandez et al., 2010). In the case of ungerminated spores which have an exceptionally robust cell wall (Bushnell and Roelfs, 1984), the addition of acid washed sand prior to grinding in liquid nitrogen has been used successfully in this laboratory for rapid disruption of this samples before DNA extraction. Proteins obtained from spores using this method were resolved on SDS-PAGE gels and compared to protein extractions from spores disrupted by very extensive grinding in liquid nitrogen without sand. Given that the number of proteins bands and their apparent intensities (reflecting the relative abundance of individual proteins) were highly comparable, sand was omitted from subsequent protocols. Nonetheless, the addition of sand is useful to speed up the procedure in cases where several spore samples need to be processed. In the case of germinated spores, water penetrates the spore during incubation (the swelling process (Griffin, 1995)), softening the hard, dry tissue inside, which facilitates their disruption by simple grinding in liquid nitrogen. The low number of FACS-purified haustoria made it difficult to apply mechanical methods to disrupt
these cells. Given that it was unknown which method would allow optimal disruption, two chemical methods used widely to disrupt low-abundance or single cell samples were tested: a surfactant (SDS) and a chaotropic agent (urea) (Walker, 2005, Brown and Audet, 2008, Petriz et al., 2012). SDS is incorporated into the cell membrane, solubilizing lipids and proteins in the membrane, creating pores within the membrane and leading to full cell lysis (Brown and Audet, 2008). Urea, similar to surfactants, breaks non-covalent interactions facilitating cell lysis, protein denaturation and protein solubilization (Petriz et al., 2012). The protein profiles from haustorial cells could not be compared on SDS-PAGE gels due to the limited material available; instead, proteins from both methods were characterized directly by LC-MS. Although both methods allowed peptide detection, method B which does not include a protein precipitation step gave better results.

Proteins from germinated spores were first submitted directly to LC-MS analysis. However, as the complexity of the sample was unknown, the effect of an OFFGEL fractionation step on the resultant MS data was assessed. This allowed the identification of about 50% more peptides, and nearly 35% additional proteins than the procedure lacking the pre-fractionation step. These percentages are comparable to similar studies which have shown the benefits of fractioning complex biological samples before proteomics analyses (Hubner et al., 2008, O’Cualain et al., 2010), and justifies inclusion of the method. Following this approach, approximately 2700 proteins were detected from ungerminated spores. This is a superior result to published studies using 2-DE/MALDI-TOF MS/MS (Noir et al., 2008, Luster et al., 2010) or gel-free Multidimensional Protein Identification Technology (MudPIT or LC/LC–MS/MS) (Cooper et al., 2007) that identified 100-500 proteins from the spores of other biotrophic fungi. Interestingly, approximately 700 and 200 proteins were exclusively detected in ungerminated and germinated Pst spores respectively. This change in protein content could reflect the significant metabolic shift during the breaking of dormancy and the initial stages of germination. Unfortunately, the time available did not allow a detailed study of the differences in protein content in the two spore stages, or annotation of the protein functions or metabolic pathways represented. Previous comparative studies on dormant and germinated spores of U. appendiculatus showed that after germination, there were only small changes in the amounts of proteins involved in glycolysis, acetyl Co-A metabolism, citric acid cycle, ATP-coupled proton transport and gluconeogenesis (Cooper et al., 2007). Also, more histone proteins accumulated, indicating a possible reorganization of the nuclei prior to appressorium formation (Cooper et al., 2007). The high coverage of spore proteins obtained here for Pst has the potential to improve the understanding of those important early changes. This could include processes such as the switch to become metabolically active, the first steps in development of the pathogenic state, and even
mechanisms involved in the avoidance of the first line of plant defence responses. Proteins extracted from purified haustoria did not show the same complexity as spores and did not require a fractionation step before LC-MS. Nevertheless, to the best of my knowledge this is the first proteomic study derived from FACS-purified haustoria, proving the utility of the haustoria purification method for proteomics studies. Similar to the spore data, the 190 proteins detected in haustoria are currently under analysis to determine their functional classification and implications for the metabolic state.

The study of candidate effectors from the proteomic perspective was important because it helps to validate the existence of these genes that were largely only ever predicted in silico. Nonetheless, the number of candidate effector proteins detected using proteomics is considerably smaller than what can be achieved with transcriptomics, indicating that proteomics still does not provide total knowledge as do other nucleic acid-based “omics” techniques. A preliminary search of peptides matching the predicted set of effector gene candidates showed that 25 of these genes could be detected at the proteomic level in haustoria. Proteomic studies on haustoria-enriched samples of other fungal biotrophic pathogens such as B. graminis (Godfrey et al., 2009, Bindschedler et al., 2009) and Puccinia triticina (Song et al., 2011), detected very few potential effector proteins from the original predicted gene sets (Xu et al., 2011, Spanu et al., 2010). It is interesting that despite the transcriptomic data that suggests that a large number of these genes are massively expressed in haustoria, only a small fraction of these could actually be found in the protein form. One possibility could be that once synthesized, these proteins are immediately secreted, so the amount of the actual protein is too low to be detected in haustoria. It is also possible that effector synthesis is stopped as soon as the haustoria are released from infected tissue, and consequently their protein levels decay rapidly during the haustoria isolation process. Additionally, proteomics is far less sensitive than transcriptomics as it deals with far more complex biological samples and has a number of technical limitations as mentioned above. This, combined with low amounts of starting material as is the case for FACS isolated haustoria, affects the efficiency of protein identification. Nonetheless, the validation of these genes from a proteomic approach is a technical achievement and makes them interesting candidates for future functional characterization.

A larger number of proteins corresponding to effector candidate genes were detected in the germinated and ungerminated spore samples. Interestingly, 24/140 of these proteins had partial annotation as glycoside hydrolases, peptidases or detoxifying enzymes, typical of secretomes of hemibiotrophs such as M. oryzae (Kim et al., 2013, Wang et al., 2011c). These proteins have been
proposed to have important functions in the early points of infection, facilitating penetration and detoxifying the harsh leaf apoplastic environment (Kim et al., 2013). Further investigation is necessary to determine if the Pst proteins could play similar roles in establishment of biotrophy. The identification of candidate effector proteins in spore tissue was not entirely unexpected despite the dogma that effector genes should be expressed predominantly in haustoria, with some expressed from infectious hyphae. This is because the expression profiles suggested that at least 27% of the total predicted effector gene set are expressed to some extent in germinated spores (Garnica et al., 2013). However, comparisons of transcriptomic and proteomic datasets from germinated spores showed that for some of these genes, the mRNA presence was uncorrelated with that of protein levels. In particular, for some detected proteins the expression of the corresponding gene was not detectable. Furthermore, the largest number of candidate effector proteins detected by proteomics was in ungerminated spores, a pathogenic stage that is metabolically inactive (Griffin, 1995). Some candidate effector genes detected at the protein level in ungerminated spores with massive expression in haustoria were tested by RT-PCR to determine if mRNA expression was indicative of protein accumulation in a given pathogenic stage. Four of these genes were not detected at the mRNA level in ungerminated spores, suggesting that their protein products were preformed before deposition into spores at some point during biogenesis. Why this happens, the role of these molecules in a place other than haustoria, where these proteins were originally synthesized, and how they are trafficked into newly formed spores, are some of the questions that arise from these findings. Although cross-contamination or inaccurate protein identification leading to experimental artifacts are improbable explanations for my observations, additional research including more proteomics replicates, transcriptomics data from ungerminated spores, immunolocalization analysis and RNA-in situ hybridizations are necessary to corroborate and better understand the results obtained in this part of my study.

It is generally believed that biotrophic fungi do not obtain significant nutrients from the host until the haustorium is established; therefore it would be expected that the spores contain substantial energy reserves to sustain the early stages of germination, elongation and establishment of pathogenic structures before successful infection (Griffin, 1995, Cooper et al., 2007). Preformed mRNAs are also deposited into ungerminated spores, to support early protein synthesis prior to the initiation of RNA synthesis (Linz and Orlowski, 1982, Griffin, 1995). So far, effector proteins in biotrophic rust fungi have not been considered as essential spore components. Numerous studies on flax rust have shown that host R genes which recognize haustorially-expressed secreted proteins encode cytosolic proteins (Lawrence et al., 2007). Screening of an haustorium-specific cDNA library...
led to the identification of three secreted flax rust effector proteins, which are recognized in the host cytoplasm (Catanzariti et al., 2006). Immunolocalization studies of the flax rust effector AvrM revealed this protein inside the host cells containing haustoria (Rafiqi et al., 2010), similar to what has been observed for the bean rust effector RTP1 (Kemen et al., 2005). These lines of evidence have been interpreted in a generally accepted model in which the synthesis and secretion of effector proteins in rust fungi is a haustorial process. However, since effectors are generally associated with positive roles in establishing infection (Alfano & Collmer 2004), the secretion of effectors in pre-haustorial stages would be advantageous for establishment of and colonization by the pathogen. Ungerminated and germinated spores need to overcome dormancy and penetrate the leaf while evading plant defence responses, in which case the presence and secretion of effector proteins in these early stages seems essential. Gene expression analyses of the subepidermal uredinium (the structure from which newly formed spores are produced) containing uredospores and sporogenous hyphae of the poplar rust *M. larici-populina* (Hacquard et al., 2010) and the soybean rust *Phakopsora pachyrhizi* (Tremblay et al., 2008), have been performed by laser microdissection of the tissues followed by whole-genome exon oligoarrays and EST sequencing respectively. Interestingly, both studies found the expression of some genes encoding predicted secreted proteins in the uredinium, although they were not directly linked with the formation of new spores; instead it was suggested that the expression of those genes was for maintenance of the biotrophic phase at the uredinial stage. Based on the proteomic and gene expression data obtained in this project, it would be plausible to hypothesize that some effector proteins could have destinations other than the host cell cytoplasm, and more importantly, that their synthesis and deployment could occur at different spatiotemporal sites. However, many critical gaps remain for future studies to elucidate.

The recently assembled genome of *Pst* 104 E137 A- (the strain used in this project) (W. Jackson, The Australian National University, personal communication), and the comprehensive set of genes predicted from this assembly and all of the transcriptomic data generated in this PhD project, represents a great opportunity to exploit the proteomic data generated for this pathogen further. The use of the combined gene sets as reference database to match MS spectra will help to detect additional proteins of *Pst* during the development from ungerminated spores to the formation of haustoria. The high-throughput proteomic data from spore samples potentially will help to discover some of the key proteins during the transition from a dormant spore to a metabolically active state, which could be of vital importance in the initiation of pathogenicity. Furthermore, the genome sequence will allow the prediction of additional effector gene candidates which would be transcriptomically and proteomically characterized based on the data generated and described in
this study. Effector gene candidates for which the mRNA localization does not correlate with their cognate protein presence will be detected, allowing further assessment of their significance to pathogenicity.
Chapter 6
General discussion

6.1. PERSPECTIVES

RNA plays a key role in determining what actually gets built in a cell. Gene-expression analyses, based on measurements of RNA levels, can reveal why two species, tissues, or cells under different conditions, differ markedly in form and function. For years, gene expression analyses on non-model organisms whose genome sequences were not available were extremely limited. Rather than ask which genes under a particular timepoint or environmental circumstance were turned on, frequently the only option was to look for genes already identified in model organisms (Maxmen, 2013). With the emergence of high-throughput next generation sequencing of RNA (RNA-seq), this limitation has been overcome. RNA-seq allows one to discover and profile gene activity in organisms that do not have available genomic resources, by building a transcriptome which contains all of the RNA molecules produced by an organism or tissue in a specific time point or stage. Transcriptomes have two characteristics that make them interesting options for studying non-model organisms: they are usually cheaper to build than genomes, and their study can eventually be more reliable than sequencing a genome, especially in organisms with repetitive genomes that can only be assembled
accurately with great difficulty (Maxmen, 2013). *Pst* seemed to be an ideal candidate to be studied through RNA-seq, as it is far from a model organism, with a complex dikaryotic genome that is highly heterozygous and almost half of it composed of repetitive elements (Zheng et al., 2013). Thus, encouraged by the possibility to explore the molecular side of this important pathogen through the study of its transcriptome, this PhD project was started. But, which specific aspects of this pathogen should be studied to gain insight into the molecular factors that control the infection, colonization and maintenance of a close parasitic relationship with the host?

Since 1984, after the cloning of the first plant pathogen avirulence protein in a bacterial pathosystem (Staskawicz et al., 1984), many studies have produced evidence to support the idea that plant pathogens including bacteria, fungi, oomycetes and nematodes, secrete virulence proteins known as effectors into plant cells to favour parasitism (reviewed by Alfano 2009). While it is widely accepted that effectors are deployed to suppress host defences, it is also likely that they have other roles, such as reprogramming of the host’s transcriptional and metabolic programs to the pathogen’s benefit (Bozkurt et al., 2012). Therefore, it is now well accepted that studying effector repertoires and their functions should give clues on how a biotrophic fungal plant pathogen can infect its host, and how the long-lasting parasitic plant-pathogen relationship is maintained. The status of molecular research on stripe rust was scarce at the time this project started. The pathogen had not previously been studied using any of the RNA-seq sequencing platforms and its genome sequencing project at the Broad Institute in the United States was still underway. EST studies (Zhang et al., 2008, Yin et al., 2009) and full-length cDNA libraries (Wang et al., 2009b, Ling et al., 2007) had been the main approaches to study *Pst* in previous years. However, the low number of sequenced clones and/or the fact that these approaches do not allow precise quantification of gene expression, limited the insight that these studies could provide. Candidate effector genes were identified in only one study, in which 6 out of the 15 candidates were induced during infection (Yin et al., 2009). Thus, with this scenario and the availability of high-throughput next generation sequencing platforms, I undertook a large-scale transcriptomic study of *Pst* to identify its effector genes and other genetic factors potentially associated with its mechanisms of pathogenicity. Although the deep exploration of the *Pst* transcriptome was the focus of this study, genomics (not included in this thesis) and proteomics also made up part of the strategy to investigate this pathogen. Strikingly, the genomic and transcriptomic landscape for *Pst* has changed dramatically during the past three years. This study shows how the sequencing of the *Pst* transcriptome through the combination of 454 and Illumina platforms, and its assembly in the absence of the genome sequence, can be used to identify candidate effector genes. Furthermore, the comparison and annotation of the transcriptomes of
haustoria and germinated spores revealed a molecular picture of the pathogen’s metabolic states during infection, and allowed the generation of an expression profile for the repertoire of effector gene candidates (Garnica et al., 2013). A catalogue of transcripts encoding non-effector genes with potential key roles in pathogenicity is now available for Pst (Garnica et al., 2013), a valuable resource for the scientific community. Through the development of a new technique to isolate haustoria of high purity (Garnica and Rathjen, 2014), it was possible to take the molecular analysis a step further by investigating the proteomic profile of this pathogenic structure (Garnica et al., in preparation). Proteomic data from spores and haustoria allowed the validation of a subset of effector candidates at the protein level. The preliminary analysis of proteomic data from spores gave important insights into a possible disconnect between effector gene expression and tissue localisation of the cognate proteins, a phenomenon that has not yet been described in other rust fungi. In parallel to this study, draft genomes of Pst strains from North America and United Kingdom were assembled (Cantu et al., 2011, Cantu et al., 2013), as well as their repertoires of candidate effector genes (Cantu et al., 2013). All together shows the impact that the next generation sequencing platforms have had on the research status of stripe rust within a remarkably short period of time.

6.2 FUTURE DIRECTIONS

The prediction of 437 effector gene candidates from the haustorial transcriptome, the analysis of their expression profiles, and the proteomic evidence for the existence of a subset of them, opens a series of possibilities to investigate Pst from the pathogenic and evolutionary points of view. Evolutionary analyses of fungal effector genes have shown that they can undergo multiple lineage-specific expansions and losses, and that they evolve putative new functions involved in adaptation to stresses and new ecological niches (Stergiopoulos et al., 2012). To date, there are not many descriptions of effector evolution in biotrophic fungal pathogens since their effector repertoires have only started to be uncovered very recently. Nevertheless, in pathogens such as Blumeria graminis, the analysis of its effector gene repertoire and their clustering into families led to the hypothesis that these families are evolving through diversifying or purifying selection between paralogs (Pedersen et al., 2012). The Pst isolate used in this study, 104 E137 A-, was the first incursion of stripe rust in Australia detected in 1979. Since then, it has evolved in isolation and in the absence of sexual recombination, resulting in the emergence of more than 19 new pathotypes over a period of more than 30 years (Wellings, 2007, Wellings, 2010a). A thorough collection of these strains is maintained at the Plant Breeding Institute of the University of Sydney at Cobbitty, NSW. This archive represents a unique opportunity to study the effect of evolutionary changes on effector
genes and their association with changes in pathogenicity, virulence and aggressiveness of the pathogen. Furthermore, the recent assembly of a high-quality draft genome sequence of the same founder isolate (W. Jackson, The Australian National University, personal communication), offers additional options for the study of the effector candidates, wherein the genomic location and genomic context (i.e. proximity to transposable elements, effector clusters, gene duplication) could play roles in their functional evolution.

The analyses of the effector candidate sequences and their predicted proteins did not provide enough clues to decipher their potential roles. They do not have similarity to genes with known roles, do not have conserved amino acid sequence motifs as found in B. graminis (Godfrey et al., 2010) or oomycete (Morgan and Kamoun, 2007) effectors, and most of them could not be classified into families or clusters with common structural features. Thus, what do the effector gene candidates or the proteins they encode have in common such that subsets of them can be coordinately expressed, and presumably transported by the same mechanism to the plant-pathogen interface? Which aspects other than those mentioned above should be studied on the set of effector gene candidates to provide clues? Two very recent studies offer interesting ideas. Marin et al (2013) examined published structural data of bacterial effectors and concluded that these proteins are highly enriched in long flexible segments that have no ordered secondary structure under physiological conditions, but can fold in a stimulus-dependent manner. The authors suggested that the high incidence of intrinsic disorder in effectors is a result of positive evolutionary selection, highlighting the dynamic nature of these proteins. They postulated that the disordered state is essential for central aspects of effector biology such as (1) effector translocation, (2) evasion of the innate immune system, and (3) host function mimicry (Marin et al., 2013). Examples of disordered regions in fungal and oomycete effectors have been reported recently in the effector PWL2 of M. oryzae, (Schneider et al., 2010) and in the effector AVR3a of P. infestans (Yaeno et al., 2011), suggesting this idea may also be applicable to eukaryotic pathogens including rusts.

The second study on the hemibiotroph M. oryzae provided evidence of two separate and specialized effector secretion systems during biotrophic invasion of rice (Giraldo et al., 2013). Apoplastic effectors seem to use the conventional ER-to-Golgi secretory process to reach the extracellular space, while effectors destined to be transported to the host cell cytoplasm appear to require the exocytosis complex for delivery. Thus, if effector targeting to distinct host compartments requires separate secretory processes, are there specialised structural characteristics of the effector proteins that tags them to one secretion system or the other? Do these dual-mode secretion systems exist in
rusts? Interestingly, the preliminary analysis of the proteomic data described in Chapter 5 indicated that some of the proteins encoded by effector gene candidates of *Pst* are deposited into newly produced spores after being synthesized elsewhere, for example in the haustorium. Could this suggest that their delayed secretion, separate from the site of synthesis, indicates the existence of a regulated secretion process in stripe rust? For this to occur, at least three conditions should be fulfilled: 1. These effector proteins are trafficked from the site of synthesis by a specialised mechanism 2. These effector proteins are presumably held in some type of membrane bound body in spores, and 3. The secretion of these effectors from spores must be highly regulated, reminiscent of the *Magnaporthe* situation. Current efforts in this laboratory are aimed at mining the current and newly-generated data to identify the best subset of candidate effector proteins which could undergo selective sorting to different secretion systems. The study of this subset of effectors may give insight into these hypotheses, perhaps through the use of heterologous fungal expression systems such as *M. oryzae* and yeasts.

The catalogue of *Pst* transcripts corresponding to genes which belong to major metabolic pathways, their expression levels and their contextualization into pathological scenarios was a major component of this study. From here, the identification, cloning and characterization of genes of interest will be facilitated by the resources generated in this project. However, since transformation methods have not yet been established for *Pst*, classical molecular methods to characterize genes such as gene inactivation by knockout, cannot be applied to this organism. Thus, what is the real value of this catalogue and how it can be further exploited? Statistical data from the protein sequence database UniProtKB/Swiss-Prot released in December 2013 shows that approximately 11,000 new entries were added per year since 2010. Other databases such as NCBI-nt and NCBI-nr (widely used in this study) are growing at an even faster pace. Taking into account that only a percentage of the whole transcriptome of *Pst* could be annotated in this study, mainly through the use of similarity searches against publicly available databases, and that the number of descriptions of new gene functions is growing rapidly, the content of the *Pst* transcriptome has enormous potential to be further understood through future annotation processes. Additionally, in this laboratory, a new method to transform fungal rusts is being developed and preliminary assays have shown positive results in flax rust (B. Conlan, The Australian National University, personal communication). The transformation of an avirulence gene into a virulent rust strain induced HR in plants carrying the corresponding resistance gene, providing an effective assay system (B. Conlan, The Australian National University, personal communication). This new method is currently being tested in *Pst*, however as no *Avr* genes have yet been identified in this pathogen, alternative
screening methods for successful transformation are required. The catalogue of annotated transcripts generated in this study has already been utilised in this laboratory as a source of specific gene candidates that, when silenced through the new transformation method, may result in an easy to differentiate phenotype. I anticipate that once the transformation method for Pst is established, the gene catalogue will be an invaluable source of candidates which can then be characterized to determine roles in pathogenicity. A third way to exploit the transcriptomic data is through the comparison with proteomic data. In Chapter 5, proteomic data from germinated spores did not always seem to correspond with the transcriptomic profile, suggesting that the transcription, protein synthesis and final destinations of several proteins (including effector candidates) might be separated spatiotemporally. The transcriptomic and proteomic datasets described here, together with the new transcriptomic data from ungerminated spores, will increase our understanding of the mechanisms used by the fungal thallus to prepare the material necessary for future infections. Future studies on specific genes that exhibit uncoupled transcription and protein synthesis to study their spatiotemporal dynamics include immunolocalization analyses, in situ RNA hybridizations, and eventually fluorescent tagging (through the new transformation method).

Another tool developed during this study was the method to ultrapurify haustoria. None of the purification methods developed so far can eliminate plant derived contamination, which is not a significant problem for transcriptomic analysis because only plant mRNAs are polyadenylated, but does have a negative impact in proteomics studies. In this work I was able to show that haustoria isolated by this method remain viable and are suitable input material for proteomic analysis. Therefore, this method represents a technological advance for the study of haustoria-forming fungi, since it could easily be adapted to isolate haustoria from other fungal species. In addition, highly purified haustoria used for transcriptomic analysis could be useful to enrich fungal mRNA at the early stages of infection, when the fungal biomass is extremely low. In a recent transcriptome study of barley powdery mildew during early pathogenesis on immunocompromised Arabidopsis, a total of two billion reads were sequenced, from which more than 90% were assigned to the Arabidopsis genome and only 1.79% could be aligned to the fungal genome (Hacquard et al., 2013). The cost of generating such an amount of data can be significant and hence not affordable to all laboratories. Thus, the enrichment of haustoria using the method presented here could be a valuable alternative to facilitate early stage-specific gene expression profiling. Highly purified haustoria isolated in an intact state and in sufficient quantity could also be used to raise specific antibodies against this structure, which could have applications on developmental and structural studies of the pathogen.
Lastly, pure and viable haustoria could be useful for investigating their cellular function thorough metabolic studies (nutrient uptake and metabolism), and effector secretion studies.

The proteomic data from different pathogenic stages of *Pst* was mainly thought of as a tool to validate a significant portion of the transcriptomic and genomic predictions. However, the interesting clues obtained from the proteomics data about the dynamic expression of effector and non-effector genes (discussed above) highlights the value of this technology. While genomics provides a static snapshot of the diversity of genes in a cell and transcriptomics gives a picture of the gene activity at a particular point in time, they do not necessarily reflect the actual proteomic state of a cell (Persidis, 1998). Further exploration of the *Pst* proteomics data will provide a more accurate way to understand how and when genome-encoded events (i.e. protein translation) occur and what relationship non-genome-encoded events (i.e. protein turnover, post-translational modifications and proteolytic processing) have in specifying a particular physiological state. Integration of the mRNA and protein expression data into a common framework should provide more details of the complex pathogenic processes of stripe rust.

Overall, this work pioneered the application and integration of high-throughput “omics” technologies to an important agricultural pathogen which, because of its biotrophic nature, had been scarcely investigated from a molecular perspective. Stripe rust soon will have comparable genomic, transcriptomic and proteomic resources to the other two wheat rusts, which will allow the comparison between these three pathogens with very similar life styles but which are causal agents of different diseases. The results presented in this thesis represent the initial steps towards the understanding of the pathogenicity factors and the metabolically essential genes which enable the pathogen to infect, colonise and parasite wheat plants. Genomic (W. Jackson, The Australian National University, personal communication) (Cantu et al., 2011, Cantu et al., 2013), transcriptomic (Garnica et al., 2013) and proteomic (Garnica et al., unpublished) data from *Pst* will shortly be fully available to the scientific community. All of these resources will enable the generation of meaningful hypotheses that will open new avenues in biotrophy research.
Appendices
Appendix 1.

List of primers to validate candidate effector gene expression by non-quantitative RT-PCR.

<table>
<thead>
<tr>
<th>Primer ID</th>
<th>Sequence 5’-3’</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pstv_13398-F</td>
<td>caccatgGCACACAGTGAATGTTGAGG</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_13398-R</td>
<td>TGCCGCTTGGGGGCTGGTCC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_3046-F</td>
<td>caccatgCAAAAATCTCGGAACCCTTACAC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_3046-R</td>
<td>AGCATCGGGATATCTTGGTAATCTC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_3060-F</td>
<td>caccatgATGCGATGGAATGTGGTGTTG</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_3060-R</td>
<td>TCGAAAAAGAAGATTAGCATGAAC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_3161-1-F</td>
<td>caccatgCAGATGTTGGAAGCGTTG</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_3161-1-R</td>
<td>AGCATCGGGATATCTTGGTAATCTC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_3164-2-F</td>
<td>caccatgCAAAATCTCGGAACCCTTACAC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_3164-2-R</td>
<td>AGCATCGGGATATCTTGGTAATCTC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_3271-F</td>
<td>caccatgTCTCCCATCGGTGACCTTGC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_3271-R</td>
<td>TCGAAAAAGAAGATTAGCATGAAC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_4624-1-F</td>
<td>caccatgGCTTCCGATTCGAATCCTTCC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_4624-1-R</td>
<td>CGGAGCCCCGATGACCATG</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_4784-3-F</td>
<td>caccatgCCTGAAAAATTCATCTCTGAGC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_4784-3-R</td>
<td>AACGACGCGGGCAACATAGATCGC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_4790-F</td>
<td>caccatgATGGAGGCTGGCAAAGCAC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_4790-R</td>
<td>GATATGATCTTTCTTGATTTGGGC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_4804-1-F</td>
<td>caccatgGCGAGTCCGCTTTCGATTCCG</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_4804-1-R</td>
<td>CTCATTGGCAAAGGAGGATCTC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_4823-F</td>
<td>caccatgAGATATGGGAGGGAGATCTC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_4823-R</td>
<td>CTCATTGGCAAAGGAGGATCTC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_4852-F</td>
<td>caccatgGCACACAGTGAATGTTGAGG</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_4852-R</td>
<td>TTTCTGAAAAAGGCCACGCGC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_4853-F</td>
<td>caccatgGAAGACTGCTTGGTCTTTCAGGCC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_4853-R</td>
<td>CCAGTCAAGGGGACCTTCG</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_4902-1-F</td>
<td>caccatgCAAGTCGCGCTTTCGAGTGG</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_4902-1-R</td>
<td>GGCAGATTGGCTCTTTTGTTCG</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_5075-F</td>
<td>caccatgGAAACGATGTGGCTGCCAGC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_5075-R</td>
<td>AAGCCCAAACTCATGAACCTC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_5088-1-F</td>
<td>caccatgAGAACTGGCAACAAAACACCAC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_5088-1-R</td>
<td>TGGTTTCTTGCTGCGGCTTCC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_5365-F</td>
<td>caccatgGTGCTCATGTTGAAACACCAC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_5365-R</td>
<td>GGCAGATTGGCTCTTTTGTTCG</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_8948-F</td>
<td>caccatgAAACGACGCTCCGGCCCACC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_8948-R</td>
<td>TTAGAGAAGAGAGAGAGCTCCAG</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_9199-F</td>
<td>caccatgGACGAGCCTGCGGAGAGAATC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_9199-R</td>
<td>AAACGCGGATTTCTGCTTCC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_12358-F</td>
<td>caccatgTATCACAGTCTTCGTTGCTAAAG</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_12358-R</td>
<td>TAGTTTGCAGCTTTTCAACATTGAAC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Accession</td>
<td>Primer Set</td>
<td>Sequence</td>
</tr>
<tr>
<td>-----------</td>
<td>------------</td>
<td>----------</td>
</tr>
<tr>
<td>Pstv_12461-F</td>
<td>caccatgACTGGACCCCCACATGAGAAAC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_12461-R</td>
<td>CTCCTTTTTCCCTGTTGAAGTCTTC</td>
<td></td>
</tr>
<tr>
<td>Pstv_12651-F</td>
<td>caccatgATCTTACTTCACAAACCGGTTG</td>
<td>RT-PCR and TOPO cloning, STOP included</td>
</tr>
<tr>
<td>Pstv_12651-R</td>
<td>CATGTTCACTGTAAGCTGAGTGC</td>
<td></td>
</tr>
<tr>
<td>Pstv_12657-F</td>
<td>caccatgACCCCGGCCTCGATACCATG</td>
<td>RT-PCR and TOPO cloning, STOP included</td>
</tr>
<tr>
<td>Pstv_12657-R</td>
<td>CAAACAAAGAGAGAAATCTTC</td>
<td></td>
</tr>
<tr>
<td>Pstv_12696-F</td>
<td>caccatgGAATTGGCTCCAGTAGGGGT</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_12696-R</td>
<td>CATGTTCACGTACTGGAAGTGGTC</td>
<td></td>
</tr>
<tr>
<td>Pstv_12697-F</td>
<td>caccatgACCCCGGCCCTCGATTCCATG</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_12697-R</td>
<td>CAAACAACAGCAACAGAATCTCC</td>
<td></td>
</tr>
<tr>
<td>Pstv_12818-F</td>
<td>caccatgTCACCCGTCGCTCTCGAACCC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_12818-R</td>
<td>CGGAATTGTGCACCCTGG</td>
<td></td>
</tr>
<tr>
<td>Pstv_13351-F</td>
<td>caccatgTACCCGAACTTCATCGCTAAAG</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_13351-R</td>
<td>AAGGTGGGAGTAAATCACCTTTTC</td>
<td></td>
</tr>
<tr>
<td>Pstv_13424-F</td>
<td>caccatgACATGCGACAAAGAGATAGC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_13424-R</td>
<td>GCCAGAGTAGTTTTTGTCAGTGC</td>
<td></td>
</tr>
<tr>
<td>Pstv_15878-F</td>
<td>caccatgATGGACCACCTTGGCCCAC</td>
<td>RT-PCR and TOPO cloning, STOP included</td>
</tr>
<tr>
<td>Pstv_15878-R</td>
<td>CTACTGCTCATTCATGAAGATTTTTAC</td>
<td></td>
</tr>
<tr>
<td>Pstv_15884-F</td>
<td>caccatgTACCCGCTCAATGAGCAATTG</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_15884-R</td>
<td>GTTTCGCAAAGTGAAGGCTACCG</td>
<td></td>
</tr>
<tr>
<td>Pstv_15910-F</td>
<td>caccatgGCAGTAGGCGATCCAGGGGGTC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_15910-R</td>
<td>AGAGTGAGGGGGAAAAGTGGG</td>
<td></td>
</tr>
<tr>
<td>Pstv_16344-F</td>
<td>caccatgATCATTAACGTACGCACGACAAG</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_16344-R</td>
<td>GCCAGAGTAGTTTTTGTCAGTGC</td>
<td></td>
</tr>
<tr>
<td>Pstv_3174-F</td>
<td>caccatgACTGGACCCCCACATGAGAAAC</td>
<td>RT-PCR and TOPO cloning, STOP not included</td>
</tr>
<tr>
<td>Pstv_3174-R</td>
<td>CAAATGATCATAATAGTTGAAG</td>
<td></td>
</tr>
<tr>
<td>Pstv_3492-F</td>
<td>CAGGATCTAGTGCAGGAGCAAGAAG</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>Pstv_3492-R</td>
<td>GCCAGAGTAGTTTTTGTCAGTGC</td>
<td></td>
</tr>
<tr>
<td>Pstv_4855-F</td>
<td>caccatgACTGGACCCCCACATGAGAAAC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>Pstv_4855-R</td>
<td>CACACCTCAGTACAGGAATCG</td>
<td></td>
</tr>
<tr>
<td>Pstv_9070-F</td>
<td>caccatgACTGGACCCCCACATGAGAAAC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>Pstv_9070-R</td>
<td>CACACCTCAGTACAGGAATCG</td>
<td></td>
</tr>
<tr>
<td>Pstv_12603-F</td>
<td>caccatgACTGGACCCCCACATGAGAAAC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>Pstv_12603-R</td>
<td>CACACCTCAGTACAGGAATCG</td>
<td></td>
</tr>
<tr>
<td>Pstv_12661-R</td>
<td>TCTACCAATCTTCGGAGGACTTCG</td>
<td></td>
</tr>
<tr>
<td>Pstn_12695-F</td>
<td>GTGATGATTGAGGTCATGCACCATAAC</td>
<td></td>
</tr>
<tr>
<td>Pstn_12695-R</td>
<td>TATGTTTTTCTCCCCATTAACAGCC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>Pstn_13443-1-F</td>
<td>ATGGAACTCCCAAAGTGGTTAGG</td>
<td></td>
</tr>
<tr>
<td>Pstn_13443-1-R</td>
<td>GATGGCCAAAGCCTAAATATCAG</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>Pstn_15540-F</td>
<td>TGAAGATTTCAGGAAGACTGAGG</td>
<td></td>
</tr>
<tr>
<td>Pstn_15540-R</td>
<td>ATCATCTTTTCTCCCCATTAACAGCC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>Pstn_15542-F</td>
<td>GTGATGATTGAGGTCATGCACCATAAC</td>
<td></td>
</tr>
<tr>
<td>Pstn_15542-R</td>
<td>TATGTTTTTCTCCCCATTAACAGCC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>Pstn_15556-1-F</td>
<td>ATGGAACTCCCCAAAGTGGTTAGG</td>
<td></td>
</tr>
<tr>
<td>Pstn_15556-1-R</td>
<td>GATCATCTTTTCTCCCCATTAACAGCC</td>
<td></td>
</tr>
<tr>
<td>Pstn_15848-F</td>
<td>GCATATCCACAGGTCATTGAAAAAG</td>
<td></td>
</tr>
<tr>
<td>Pstn_15848-R</td>
<td>CAACATTGGATCATGCGGAACACCAG</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>Pstn_15882-1-F</td>
<td>GCTCCCATGGTCGGGGATGTGGTCC</td>
<td></td>
</tr>
<tr>
<td>Pstn_15882-1-R</td>
<td>GGTTGGCAGCAATGGCAACGTGACC</td>
<td></td>
</tr>
<tr>
<td>Pstn_15883-3-F</td>
<td>AATGTCCCTTTTCAACACTTGCGG</td>
<td></td>
</tr>
<tr>
<td>Pstn_15883-3-R</td>
<td>GCTCCCATGGTCGGGGATGTGGTCC</td>
<td></td>
</tr>
<tr>
<td>Pstn_15883-4-F</td>
<td>GTCAACTCTCCATCCGACGATGCC</td>
<td></td>
</tr>
<tr>
<td>Pstn_15883-4-R</td>
<td>GCTCCCATGGTCGGGGATGTGGTCC</td>
<td></td>
</tr>
<tr>
<td>Pstn_15889-1-F</td>
<td>AACAACGCGTCTGAAAGCTCAGTGTG</td>
<td></td>
</tr>
<tr>
<td>Pstn_15889-1-R</td>
<td>TGAAGGTTGCATATCATCCGAAG</td>
<td></td>
</tr>
<tr>
<td>Pstn_15897-F</td>
<td>AGCCCAAACCTCTCGAGGGAGTGTG</td>
<td></td>
</tr>
<tr>
<td>Pstn_15897-R</td>
<td>AAGTCCACAGGTCATTGAAAAAG</td>
<td></td>
</tr>
<tr>
<td>Pstn_16330-F</td>
<td>TACCCAAGCTTCATCACTAAAAATTTAG</td>
<td></td>
</tr>
<tr>
<td>Pstn_16330-R</td>
<td>AAGGTGGGAGTAAATCACCTTTTCA</td>
<td></td>
</tr>
<tr>
<td>Pstn_16432-F</td>
<td>AGCCCAATGCCTCAAATACTAGAA</td>
<td></td>
</tr>
<tr>
<td>Pstn_16432-R</td>
<td>ATGGAAATGGTGACAGATTGTACAC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>Pstn_3032-F</td>
<td>CAATATGGAATGGAAGCTGGTTAG</td>
<td></td>
</tr>
<tr>
<td>Pstn_3032-R</td>
<td>GAGAAAGAAGGGCAATACTGGTG</td>
<td></td>
</tr>
<tr>
<td>Pstn_3169-F</td>
<td>AGCCCAACAATCGAGGGCTAG</td>
<td></td>
</tr>
<tr>
<td>Pstn_3169-R</td>
<td>TGAAGGTTGCATATCATCCGAAG</td>
<td></td>
</tr>
<tr>
<td>Pstn_4629-F</td>
<td>GACTCTGGGGCTCCAAAAGTCG</td>
<td></td>
</tr>
<tr>
<td>Pstn_4629-R</td>
<td>TGAAGGCTACCCATTTACCTTC</td>
<td></td>
</tr>
<tr>
<td>Pstn_4633-F</td>
<td>AGGAGGTGGTGGAGGAGAGAGAGAGAG</td>
<td></td>
</tr>
<tr>
<td>Pstn_4633-R</td>
<td>TGAAGGTTGCATATCATCCGAAG</td>
<td></td>
</tr>
<tr>
<td>Pstn_4657-F</td>
<td>AGCCCAACAATCGAGGGCTAG</td>
<td></td>
</tr>
<tr>
<td>Pstn_4657-R</td>
<td>TGAAGGTTGCATATCATCCGAAG</td>
<td></td>
</tr>
<tr>
<td>Pstn_4820-F</td>
<td>CAGCTACTCTGAGGTCATTGATG</td>
<td></td>
</tr>
<tr>
<td>Pstn_4820-R</td>
<td>GACTCTGGGGCTCCAAAAGTCG</td>
<td></td>
</tr>
<tr>
<td>Pstn_4889-F</td>
<td>CAATATGGAATGGAAGCTGGTTAG</td>
<td></td>
</tr>
<tr>
<td>Pstn_4889-R</td>
<td>GAGAAAGAAGGGCAATACTGGTG</td>
<td></td>
</tr>
<tr>
<td>Pstn_4917-F</td>
<td>ATGGAAATGGAATGGAAGCTGGTTAG</td>
<td></td>
</tr>
<tr>
<td>Pstn_4917-R</td>
<td>TTTTGGGAGTAAATCACCTTTTCA</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>Accession</td>
<td>F</td>
<td>R</td>
</tr>
<tr>
<td>-----------</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>PSTV_4962</td>
<td>AATAAGAATAAGATAATATGCGCTC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_4962</td>
<td>AGCATCTTCCTAGTTGGTAGCAAGAC</td>
<td></td>
</tr>
<tr>
<td>PSTV_7145</td>
<td>CAGGAGCATCAAGCAGCATACACAC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_7145</td>
<td>TCCTGCTGCTGCTCTGAGGACGAGGAC</td>
<td></td>
</tr>
<tr>
<td>PSTV_8952</td>
<td>ACACCGAGCATCAAGCAGCATACAC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_8952</td>
<td>AGAATAAGAATAAGATAATATGCGCTC</td>
<td></td>
</tr>
<tr>
<td>PSTV_12685</td>
<td>AGGAGCTACAGACTGAAGTGCTGGG</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_12685</td>
<td>TGAGTATCTCTCTCCCGTTTCTAAG</td>
<td></td>
</tr>
<tr>
<td>PSTV_16080</td>
<td>TGAGTATCTCTCTCCCGTTTCTAAG</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_16080</td>
<td>AGGAGCTACAGACTGAAGTGCTGGG</td>
<td></td>
</tr>
<tr>
<td>PSTV_3267</td>
<td>GAGGAGGCATCAAGCAGCATACACAC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_3267</td>
<td>GCATCAGTCGAAGCACTTAAGCC</td>
<td></td>
</tr>
<tr>
<td>PSTV_3749</td>
<td>CTGTTTCCCAGACATCAAGC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_3749</td>
<td>TCTGGAGCAAGCACACAAAGAAATGC</td>
<td></td>
</tr>
<tr>
<td>PSTV_4597</td>
<td>CAGGCTCCGCTGACCGTCAACACTC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_4597</td>
<td>TCTGGAGCAAGCACACAAAGAAATGC</td>
<td></td>
</tr>
<tr>
<td>PSTV_4922</td>
<td>GCCCCGGTATCCAGGTGGTGTTGC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_4922</td>
<td>ATCCCCATCAATGTTTCTGGGACTG</td>
<td></td>
</tr>
<tr>
<td>PSTV_5004</td>
<td>ATCCCCATCAATGTTTCTGGGACTG</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_5004</td>
<td>AGATCTGGTTCGTTGAAAACTGATC</td>
<td></td>
</tr>
<tr>
<td>PSTV_5062</td>
<td>CTCTGCCTTTTGCGACACACAC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_5062</td>
<td>AGGAGCTACAGACTGAAGTGCTGGG</td>
<td></td>
</tr>
<tr>
<td>PSTV_5115</td>
<td>ATCCCTTTTCCAGCCCAATGCTTTTC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_5115</td>
<td>CTCTGCCTTTTGCGACACACAC</td>
<td></td>
</tr>
<tr>
<td>PSTV_5442</td>
<td>CTCTGCCTTTTGCGACACACAC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_5442</td>
<td>GGAATCCGTCTGACATTGTC</td>
<td></td>
</tr>
<tr>
<td>PSTV_5121</td>
<td>GTCCCACTAGAACTGGAACAGG</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_5121</td>
<td>GCATCAGTCGAAGCACTTAAGCC</td>
<td></td>
</tr>
<tr>
<td>PSTV_6044</td>
<td>GCATCAGTCGAAGCACTTAAGCC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_6044</td>
<td>GGCATCTTTTCCAGCCAATGCTTTTC</td>
<td></td>
</tr>
<tr>
<td>PSTV_6870</td>
<td>ATGGAGGAGGAGTAGCTGCAGCTTGAAG</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_6870</td>
<td>GCAAGGAGGAGTAGCTGCAGCTTGAAG</td>
<td></td>
</tr>
<tr>
<td>PSTV_7145</td>
<td>TCTGGAGCAAGCACACAAAGAAATGC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_7145</td>
<td>GCATCAGTCGAAGCACTTAAGCC</td>
<td></td>
</tr>
<tr>
<td>PSTV_8952</td>
<td>CTCTGCCTTTTGCGACACACAC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTV_8952</td>
<td>AGGAGCTACAGACTGAAGTGCTGGG</td>
<td></td>
</tr>
<tr>
<td>PSTN_3223</td>
<td>ATCCCTTTTCCAGCCCAATGCTTTTC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTN_3223</td>
<td>CTCTGCCTTTTGCGACACACAC</td>
<td></td>
</tr>
<tr>
<td>PSTN_5062</td>
<td>ATCCCTTTTCCAGCCCAATGCTTTTC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTN_5062</td>
<td>GCATCAGTCGAAGCACTTAAGCC</td>
<td></td>
</tr>
<tr>
<td>PSTN_9213</td>
<td>ATCCCTTTTCCAGCCCAATGCTTTTC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>PSTN_9213</td>
<td>CTCTGCCTTTTGCGACACACAC</td>
<td></td>
</tr>
<tr>
<td>Primer Set</td>
<td>Sequence</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>Pstn_9213-R</td>
<td>TCCATGAACCTTGATCATCAGGTC</td>
<td></td>
</tr>
<tr>
<td>Pstv_12014-F</td>
<td>CAGGGTGGGCTCAAACAACCCAAGC</td>
<td></td>
</tr>
<tr>
<td>Pstv_12014-R</td>
<td>CTTCTCTCCATCTGAAGGTGTGC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>Pstn_12353-F</td>
<td>GAGCCTTTATGCCTGTTGTAAG</td>
<td></td>
</tr>
<tr>
<td>Pstn_12353-R</td>
<td>GCTGTCGAACCGGCTCTGCCAGG</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>Pstn_15579-F</td>
<td>GAATTTGAAATAATCAACCAATCAG</td>
<td></td>
</tr>
<tr>
<td>Pstn_15579-R</td>
<td>CTTCGGAAGAAGGATGGGCGTGCG</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>Pstv_15614-F</td>
<td>GGAACCTCCTCAATTTCTACAAGTTC</td>
<td></td>
</tr>
<tr>
<td>Pstv_15614-R</td>
<td>ACGCTTTGAGGACGGCACAACC</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>Pstv_15616-F</td>
<td>GCAGGAATATCGGACGAGGATTTTG</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>Pstv_15616-R</td>
<td>GCGATTGGGATAACGATTAGATG</td>
<td></td>
</tr>
<tr>
<td>Pstv_16438-F</td>
<td>TTCAATGTAGGTACCCGAATTTAG</td>
<td>RT-PCR, STOP not included</td>
</tr>
<tr>
<td>Pstv_16438-R</td>
<td>TTCTACGTTGACGAATTCTTGAGC</td>
<td></td>
</tr>
<tr>
<td>GaPDH v.3 F</td>
<td>TGGCTGTGCTCTCAAGTTATC</td>
<td>Specific for wheat Glyceraldehyde 3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>GaPDH v.3 R</td>
<td>CTCAGGAAGAACCCTTACCAAACA</td>
<td></td>
</tr>
<tr>
<td>PR2-F</td>
<td>CTCGAACATCGGTAACGACCAG</td>
<td>Specific for wheat PR2 (β,1-3 glucanase), accession Y18212</td>
</tr>
<tr>
<td>PR2-R</td>
<td>GCGGCGATGTACTTGATGTTC</td>
<td></td>
</tr>
<tr>
<td>PR5-F</td>
<td>ACAGCTACGCCAAGGACGAC</td>
<td>Specific for wheat PR5 (thauatin-like protein), accession AF442967</td>
</tr>
<tr>
<td>PR5-R</td>
<td>CGCGTCTAATCTAAAGGCCAG</td>
<td></td>
</tr>
<tr>
<td>TaGLP2a-F</td>
<td>AACAAAGGTGATGTTCGCTCTTC</td>
<td>Specific for wheat TaGLP2a (germin-like), accession AJ237942</td>
</tr>
<tr>
<td>TaGLP2a-R</td>
<td>GAGCCGGTCTATGGATATTTTCCT</td>
<td></td>
</tr>
<tr>
<td>Pst β-tubulin-F</td>
<td>TGATCGTGATAGGCCACCTTCTCTC</td>
<td>Specific for Pst β-tubulin, accession EF570843</td>
</tr>
<tr>
<td>Pst β-tubulin-R</td>
<td>GAACGGACAGCATGTTCTCTTCG</td>
<td></td>
</tr>
</tbody>
</table>
References


ARAFEH, R. B. 2008. Strategy to Design a Stratified Lamellar Corneal Stroma Construct. Master of Science, University of North Texas Health Science Center at Fort Worth.


CHONG, J., HARDER, D. E. & ROHRINGER, R. 186. Cytochemical studies on *Puccinia graminis* f.sp. *tritici* in a compatible wheat host. 2. Haustorium mother cell-walls at the host-cell


GAY, J. L., SALZBERG, A. & WOODS, A. M. 1987. Dynamic experimental-evidence for the plasma-membrane Atpase domain hypothesis of haustorial transport and for ionic coupling of the...
haustorium of *Erysiphe graminis* to the host-cell (*Hordeum vulgare*). *New Phytologist*, 107, 541-548.


References


HEATH, M. C., XU, H. X. & EILAM, T. 1996. Nuclear behavior of the cowpea rust fungus during the early stages of basidiospore- or urediospore-derived growth in resistant or susceptible cowpea cultivars. Phytopatology, 86, 1057-1065.


JIN, Y. 2011. Role of *Berberis* spp. as alternate hosts in generating new races of *Puccinia graminis* and *P. striiformis*. *Euphytica*, 179, 105-108.


MANNERS, J. M. & GAY, J. L. 1982. Transport, translocation and metabolism of 14C-photosynthates at the host-parasite interface of *Pisum sativum* and *Erysiphe pisi*. *New Phytologist* 91, 221-244.


References


References


Arabidopsis thaliana and Peronospora parasitica. Physiological and Molecular Plant Pathology, 65, 67-78.


WANG, J. Y., CHAMBON, C., LU, C. D., HUANG, K. W., VIVARES, C. P. & TEXIER, C. 2007. A proteomic-based approach for the characterization of some major structural proteins involved in host-
parasite relationships from the silkworm parasite *Nosema bombycis* (Microsporidia). *Proteomics*, 7, 1461-1472.


Strategies for Wheat Stripe Rust Pathogenicity Identified by Transcriptome Sequencing

Diana P. Garnica¹, Narayana M. Upadhyaya², Peter N. Dodds², John P. Rathjen¹*,

¹Research School of Biology, Australian National University, Canberra, Australian Capital Territory, Australia, ²Division of Plant Industry, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Canberra, Australian Capital Territory, Australia

Abstract

Stripe rust caused by the fungus *Puccinia striiformis* f.sp. *tritici* (*Pst*) is a major constraint to wheat production worldwide. The molecular events that underlie *Pst* pathogenicity are largely unknown. Like all rusts, *Pst* creates a specialized cellular structure within host cells called the haustorium to obtain nutrients from wheat, and to secrete pathogenicity factors called effector proteins. We purified *Pst* haustoria and used next-generation sequencing platforms to assemble the haustorial transcriptome as well as the transcriptome of germinated spores. 12,282 transcripts were assembled from 454-pyrosequencing data and used as reference for digital gene expression analysis to compare the germinated uredinospores and haustoria transcriptomes based on Illumina RNAseq data. More than 400 genes encoding secreted proteins which constitute candidate effectors were identified from the haustorial transcriptome, with two thirds of these up-regulated in this tissue compared to germinated spores. RT-PCR analysis confirmed the expression patterns of 94 effector candidates. The analysis also revealed that spores rely mainly on stored energy reserves for growth and development, while haustoria take up host nutrients for massive energy production for biosynthetic pathways and the ultimate production of spores. Together, these studies substantially increase our knowledge of potential *Pst* effectors and provide new insights into the pathogenic strategies of this important organism.

Introduction

Stripe rust is an important agricultural disease that constitutes a major challenge to wheat production worldwide. The causal agent is the fungus *Puccinia striiformis* f.sp. *tritici* (*Pst*). In Australia, wheat stripe rust causes severe losses and is controlled mainly by fungicides, with about AUD$127 million expended per annum on chemical controls [1,2]. In a worldwide context, the disease has become increasingly important in the key wheat-consuming areas of North Africa, west and central Asia [3], and China [4], and in the U.S [5] where epidemics have caused huge yield losses since 2000. The economic importance of rusts derives from three main factors; the large extent to which they can reduce grain yield, their ability to spread rapidly and reach epidemic proportions under favorable conditions, and their rapid mutation to overcome host resistance genes [6]. Thus, understanding the molecular strategies used by *Pst* to cause disease is crucial for developing more effective control methods.

Rust fungi have the most complicated life cycles of all fungi [7]. Recently, the discovery of the alternate host of stripe rust has allowed classification of this pathogen as macrocyclic because it produces five types of spores during a complete life cycle, and heterocercous, because it requires two hosts for completion of the cycle [8]. Uredinospores produced during the asexual cycle are infectious on wheat, and multiple cycles of reinfection are possible during the growing season. Rust fungi, including *Pst*, are obligate biotrophic pathogens that require living tissue to develop and reproduce. The biotrophic interaction may be sustained over many weeks, with important implications for the pathogenic lifestyle. For example, such prolonged interaction is usually associated with changes in photosynthate translocation patterns within the plant [9]. The infection site becomes a nutrient sink allowing the parasite to exploit the plant’s resources [10]. Plants constantly monitor their local environment and induce defensive responses against invasive microbes, including hypersensitive cell death that is antithetical to biotrophy. Thus the need to survive the inhospitable inner leaf environment while mining the available nutrients may recapitulate the major forces impacting the evolution of biotrophy [11].

As an obligate parasite, *Pst* has adapted to source nutrients from living mesophyll wheat cells through a sophisticated cellular structure termed the haustorium [12]. This arises as a differentiation of the infectious hyphae and invades the host cell, however it is not located within plant cell per se but is separated from the host cytoplasm by an extrahaustorial membrane which is contiguous with the plant cell plasma membrane. The haustorium is thought to be the primary organ for nutrient transfer from the host cell to the fungal vegetative body [13,14]. The space between the extrahaustorial membrane and fungal haustorial wall, called the extrahaustorial matrix, is enriched in carbohydrates with unknown function in parasitism [12]. Haustoria are not only feeding structures; they induce structural changes in the host cell.
including cytoskeletal rearrangements, nuclear migration and chromatin condensation [15], and there is evidence that they influence host cell metabolism [16,17]. Furthermore, they deliver essential virulence molecules called ‘effectors’ into the extrahaustorial matrix, several of which are subsequently translocated into host cells [18,19].

The effectors secreted by plant pathogens are believed to be essential to the parasitic lifestyle. Very little is known about the activities of plant fungal effectors, but broadly they are thought to manipulate the physiological and immune responses of host cells during infection [20]. Conversely, in certain cases the host has evolved to recognize effectors through the action of specific intracellular immune receptors. In this case, the recognized effector serves as a signal for the plant to induce defenses to block pathogen growth [21]. Historically, such effectors are termed avirulence (Avr) proteins.

The haustorium is a site of concerted host-pathogen interaction, and describing its functions is essential to understanding biotrophy. The presence of sugar transporters [14] and putative amino acid transporters [13,22] within the structure implies an important role in fungal nutrient uptake. However, some other basic questions regarding the function of the haustorium have yet to be addressed. These include the identification of mechanisms for bidirectional transport (importation of nutrients and export of effectors and other molecules), the pathogen and host proteins specifically located at the haustorium–host cell interface, and the regulatory genes which specify haustorial identity.

The availability of genomic data for biotrophic organisms has increased dramatically in the last decade providing important insights into the infection strategies of these pathogens. The genome sequences for the plant pathogens Puccinia graminis Esp triticit (Pyg, Melampsora larici-populina (Mlp) [23] and Blumeria graminis [24] share some evolutionary features of adaption to the extreme parasitic lifestyle, such as the loss of nitrate and sulfate assimilation pathways. Expressed sequence tag (EST) libraries have been generated for Pgt [25,26,27] and other Puccinia species [28,29,30], which have been useful for stage-specific expression analysis and gene prediction in genomes. However, the resolution power of these studies for completeness and comparison of transcript abundance between cell types is limited. Here, we have taken advantage of next generation-sequencing (NGS) technologies and haustorium purification methods [31] to broadly analyze gene expression in the germinated uredinospores (henceforth called “germinated spores”)and haustorial stages of Pst. The results derived from 434 and Illumina-based transcriptome sequencing allowed the prediction of haustorial secreted proteins (HSPs), a set of genes enriched in effector candidates found to be specifically- or highly-expressed in haustoria. Additionally, the comparison of the haustorial and germinated spores transcriptomes revealed fundamental metabolic differences between the two pathogenic stages.

Results and Discussion

454-pyrosequencing and Assembly

Transcriptomes corresponding to purified stripe rust haustoria and germinated spores were sequenced by single-read pyrosequencing on a 454 GS-FLX titanium platform (454). This produced 729,036 and 457,671 reads averaging 413 bp and 420 bp in length for haustoria and germinated spores, respectively. To remove contaminating wheat sequences from the haustoria data, the Pst-130 draft genome [32], a draft genome of a local Pst isolate (P-st-104E137A) (Jackson, Garnica, Foret, Rathjen and Studholme, unpublished data) and the germinated spore transcriptome generated in this study (Text S1 and Data S1) were used as references to extract fungal sequences. Since these references do not represent the complete Pst genome, the remaining unmapped reads were assembled de novo and the resulting contigs screened by BLAST search against the NCBI nucleotide and protein databases. The BLAST result was curated manually and contigs showing hits to plant genes were removed, while the remaining contigs were retained as novel transcripts not included in the draft genome assemblies. The filtered haustorial reads were then assembled de novo using CLC genomics (CLC Bio v3.9) resulting in 12,846 contigs representing the haustorial transcriptome. Assembled contigs ranged between 200 and 6,854 bp, with an average length of 704 bp (Data S2 and Figures S1, S2 and S3).

Putative Effector Candidate Genes Expressed in Haustoria

The haustoria contigs were analyzed using CLC genomics to identify all possible ORFs, which were further analysed with SignalP 3.0 [33] to predict secretion signal peptides (SP). We found 1,299 SP-encoding genes which were then filtered to include only those unique to a contig, or alternatively the largest SP-ORF within a contig. Proteins containing transmembrane domains (predicted using TMHMM 2.0 [34] or mitochondrial targeting signals (predicted using TargetP [35]) were excluded. This left 437 haustorial secreted proteins (HSPs, Data S3). The predicted mature peptide sequences (minus the SP) were searched against the NCBI non-redundant protein database using BLASTx, and were additionally analysed with BLAST2GO (B2G) [36] for the presence of conserved InterProScan protein signatures. About 60% of these genes had no similarity to known genes at the level of e-val $10^{-25}$. One hundred and five genes were similar to hypothetical proteins in Pgt, and four were similar to previously identified Pst secreted proteins [37]. Forty HSPs could be partially annotated with B2G, fifteen of which were classified as glycoside hydrolases from different families, especially family 18 (chitinase activity) and families related to plant cell wall degradation. Four HSPs were annotated as putative polysaccharide deacetylases (chitin deacetylases), and the rest were annotated as putative proteins participating in diverse cellular processes such as protein folding, proteolysis, oxidation-reduction, and regulation of transcription (Table S2, HSPs). A further 35 HSPs contained conserved protein domains such as zinc finger domain, copper/zinc binding domain, cupredoxin domain, harvin-like endoglucanase domain, NUDIX hydrolase domain, thiamatin and others. Most of these roles or domains have been identified in other rust predicted secretomes [23,30], suggesting conserved roles in Puccinales. Interestingly, Bhattacharya et. al [39] recently reported a novel effector gene from Colletotrichum truncatum (C2MUDIX) that is exclusively expressed during the late biotrophic phase and elicits the hypersensitive response in tobacco leaves, suggesting that this effector could be important for the transition from biotrophy to necrotrophy, or alternatively is recognised by the tobacco immune system. A secreted protein from Pst containing a NUDIX domain could potentially regulate redox homeostasis. Nevertheless, as observed previously for rust fungi [38], the absence of recognisable protein domains in candidate effector proteins is a common occurrence.

Genome and transcriptome sequencing of Blumeria graminis revealed that the most highly expressed candidate effector proteins contained a V/F/W/C motif 1–30 aa after the predicted SP [24]. We identified 1 to 4 copies of this sequence motif in 124 of the 437 Pst HSPs, but the motif occurred within the first 30 aa for only 43 of these, which is similar to that number expected by chance ($\sim$32). Thus similar to Pgt and Mlp [23], this motif does not seem to define a major class of effectors in Pst. Further attempts to detect
predicted HSPs (Table S2, HSPs). All tested genes showed a very close correlation between the expression pattern detected with RT-PCR, and that predicted in silico (Table S2, HSPs). Remarkably more than 85% of the HSPs were differentially expressed; 295 overexpressed in haustoria and 76 overexpressed in germinated spores. Strikingly, 40% of the HSPs overexpressed in haustoria showed no expression in germinated spores and 50% were ten or more times expressed in haustoria than in germinated spores, consistent with the idea that effectors play important roles during the biotrophic host-pathogen interaction (Figure 1B).

Functional Classification of Transcriptome Sequences

The transcripts in the reference set were categorised into functional classes using BLAST2GO to identify genes that encode proteins with known roles in cellular processes. Of this set, 4,485 transcripts could be unambiguously annotated. Using the list of the genes previously found to be differentially expressed, Fisher’s exact test was applied to find functional categories over-represented in each developmental stage. The major functional categories are shown in Figures 3 and 4. Broadly speaking, processes up-regulated in germinated spores were representative of cellular proliferation, such as cell cycle control and DNA and cell wall metabolism, whereas haustoria were more engaged in energy production and biosynthetic processes. Figure 3 refers to transcripts overexpressed in each tissue in the selected GO categories, whereas Figure 4 includes all transcripts classified under the selected GO categories and relative levels of expression were represented as colours. This is important because processes included within the same ontogenic category can have fundamentally opposed activities. We discuss the data in more detail with reference to fungal development and metabolism, below.

DNA replication and cell cycle. The two sampled tissues represent very different developmental stages, with germinated spores involved in growth and division, whereas the haustorial stage is terminally differentiated. Generally, transcripts classified into the categories of cell cycle, DNA replication and cell division were more highly represented in germinated spores (Figure 4). Genes involved in DNA replication and cell division included DNA polymerase subunits and replication factors. Cell division genes were clearly enriched in germinated spores, consistent with the idea that haustoria are non-dividing fully differentiated cells. For example, cyclins (proteins that regulate cyclin dependent kinases (CDKs) to control cell division), cyclin-dependent kinases, cohesins (protein complexes that regulate the separation of sister chromatids during cell division), septins (GTP binding proteins that provide structural support during cell division and compartmentalize parts of the cell) and genes associated with the control of mitotic phase progression were all up-regulated in germinated spores (Figure 4 and Table S6, HeatMap). Although the expression of these genes makes no implication about their function in vivo, because both positive and negative roles in control of cell cycle are possible, the observations suggest that active cell cycle regulation occurs in germinated spores. One of the most interesting transcripts found to be up-regulated in germinated spores encoded a protein similar to the cyclin-dependent kinase Cdk5 (PST79_11595). Flor-Parra et al. [43] found that the cyclin Pe12 represents a polarity- and virulence-specific regulator in Ustilago maydis, complexing with the cyclin-dependent kinase Cdk5 for sustained polar growth [43]. Although we found no Pe12 homologs in our data (nor in the Pe130 draft genome [32]), Pe12 encodes other cyclin-like proteins that might regulate Cdk5 function. Homologs of U. maydis cyclin genes B1 and B2 were also found to be highly expressed in Pe130A germinated spores, with the latter one of the most strongly up-regulated genes of this

Digital Expression Analysis of Pst Transcriptomes

To compare gene expression between haustoria and germinated spores, we used Illumina RNAseq analysis to obtain statistically robust quantitative expression data. We first assembled a transcriptome reference gene set by de novo assembly of the combined germinated spores and haustorial 454 reads, giving a set of 12,282 transcripts (Table S1, Transcripts reference set and Data S4). Three RNA samples comprising independent biological replicates were prepared for each tissue and sequenced by Illumina RNAseq analysis, giving a total of 500 million reads (Table S7). The Pst transcripts reference set was used as a template for digital differential expression analysis using the RNAseq tools from CLC genomics. Raw Illumina reads from each tissue were independently mapped against the reference set, and read counts were independently adjusted to reads per kilobase per million mapped reads (RPKM, [41]), to facilitate the comparison of transcript levels assigned to each gene between samples. RPKMs were statistically assessed using Baggerly’s test [42] (FDR corrected p-value less than 0.05) for each gene to determine differential expression between haustoria and germinated spores. A total of 4,346 transcripts were differentially expressed, revealing a clear difference between the transcriptional programs of germinated spores and haustoria (Table 1 and Figure 2).

Digital expression analysis based on Illumina sequencing also allowed us to quantify the level of expression of the 437 previously novel motifs using the MEME analysis module within CLC failed to identify other conserved sequences.

Extracellular proteins frequently contain elevated numbers of cysteine residues, which can participate in disulfide bonding to provide stability to the folded proteins in the protease-rich extracellular apoplastic space. Many known and predicted effector proteins from filamentous pathogens are small cysteine-rich proteins [17,38,40]. Analysis of the cysteine content of the predicted 437 HSP proteins revealed that 246 contained fewer than four, and 191 contained from 4–28 cysteines in the mature protein.

Validation of Haustorial Expression of Effector Candidate Genes

To validate the in silico predictions (ie to test that the genes are real, are expressed, of the expected size, and of Pst origin), 94 effector candidates were randomly chosen from the 437 HSP genes and tested for expression in silico by non-quantitative reverse transcriptase-PCR (RT-PCR) (Figure 1A). Primers were designed to amplify the full ORF excluding the SP-encoding sequence (Table S8, RT-PCR primers). RNA samples were prepared from germinated spores, isolated haustoria, wheat tissue nine days after infection (dai, maximum number of haustoria before sporulation), and uninfected wheat tissue, and subjected to RT-PCR using effector-specific primers. Primer pairs for Pst β-tubulin or wheat specific genes served as a positive controls. Seventy one showed expression in isolated haustoria and infected wheat tissue, but no detectable expression in germinated spores, 29 showed expression in all tissues where the fungus was present, and one failed to amplify. Ninety one of the tested genes were confirmed to be of fungal origin in this assay as they were amplified from genomic DNA and/or spore cDNA. The remaining three genes were also confirmed to be fungal origin as they were identified in our draft genome of this isolate (Pst104E137A) (Jackson et. al. unpublished data) but were not amplified from genomic DNA because one or both primers span intronic regions. Cloning and Sanger sequencing of 25 candidate effectors confirmed that they corresponded to the sequence predicted from our 434 assemblies.

Wheat Stripe Rust Transcriptome during Infection
Figure 1. Effector expression patterns. A. Ninety four HSPs were selected arbitrarily and their expression patterns were analysed by RT-PCR. Primers were designed to amplify the full length gene sequence minus the signal peptide-encoding region. Pst 104E137A- genomic DNA, and cDNAs from germinated spores, isolated haustoria, infected wheat leaves, and uninfected wheat tissue were used as templates. Two expression patterns were obtained: 1. Expression only during the biotrophic phase (Pstv_3161-1), and 2. Expression during early and late stages of development (Pstv_7541-1). None of the tested effector candidates were amplified from uninfected wheat leaves. B. Number of HSPs showing digital expression patterns as shown. Overexpression was evaluated using Baggerley's test [42], genes with a FDR corrected p-value less than 0.05, fold change >2 and a difference of at least 20 were considered to be significant.

doi:10.1371/journal.pone.0067150.g001
category. The *U. maydis* genes are important in cell cycle regulation [44], with mutants arrested after S phase, while overexpression of cyclin B2 generates cells with anomalous DNA content and premature entry into mitosis. In *M. oryzae*, degradation of the homologous cyclin B proteins, CYC1 and CYC2, is necessary for mitosis exit and pathogenesis [45]. The connections between cell cycle, morphogenesis and virulence seem to be important for this pathogen to colonise plant tissue, and this may also be the case for rust pathogens.

Four transcripts with similarity to septin genes (PST79_4650, PST79_4652, PST79_4126, PST79_582) were found in our data, three of which were up-regulated in germinated spores. Septins constitute a cytoskeletal structure that is conserved in eukaryotes [46]. In *Saccharomyces cerevisiae*, septins assemble as a ring that marks the cytokinetic plane throughout the budding cycle and this structure participates in different aspects of morphogenesis, such as cell polarity, localization of chitin synthesis, and the spatial regulation of septation [46]. Recently Dagdas et al. [47] showed that septin proteins in *M. oryzae* assemble in a ring-shaped structure at the base of the appressorium together with F-actin and other cytoskeletal components to provide rigidity and negative membrane curvature for protrusion of the penetration peg into the host. Although *Pst* often can form appressoria over stomata prior to leaf penetration, this is not strictly necessary as penetration can often occur without appresoria development. Thus the role of septins in leaf penetration remains to be established.

**Cell wall modification enzymes.** Cell walls are complex polysaccharide structures composed mainly of cellulose, hemicellulose and pectin in plants [48], and chitin, glucans and other

---

**Figure 2. Differential gene expression in *Pst* germinated spores and haustoria.** Venn diagram of reference transcripts set, showing the number of transcripts that did not show differential expression between the two tissues, and those that had statistically significant changes in expression between germinated spores and haustoria.

doi:10.1371/journal.pone.0067150.g002

**Figure 3. Comparative ontology analysis of transcripts with statistically significant changes in expression between haustoria and germinated spores.** Of the original transcript set, 30.2% (601) of the 1,989 haustorial-enriched genes and 47% (1,109) of the 2,357 genes up-regulated in germinated spores were annotated with B2G. Relevant biological process GO terms are shown on the Y-axis. Percentages of genes differentially expressed in each tissue belonging to the nominated categories are shown on the X-axis.

doi:10.1371/journal.pone.0067150.g003
minor components in fungi [49]. Glycoside hydrolases (GH), polysaccharide lyases, and esterases are enzymes that allow formation, remodelling or degradation of cell walls and play a fundamental role in plant-fungal pathogen systems [50]. The battery of carbohydrate-modifying enzymes derived from plant pathogenic fungi can vary enormously according with the ecological niches occupied by different species. Recent genomic studies on plant fungal pathogens have shown that the number of GH genes correlates with the type of host interaction [23,24,51,52,53]. Biotrophic fungi seem to possess a reduced number of GH enzymes, consistent with the necessity of minimizing host cell wall damage to avoid triggering plant immunity [23,24]. We found a total of 53 transcripts with similarity to GHs belonging to 18 different families that were expressed in Pst (Table S3, General metabolism). The most abundant families were GH5, GH18 and GH47, which agrees with recent findings for Pgt and Mlp, where the GH5 and GH47 families were expanded [23]. Thirteen Pst GHs were similar to enzymes implicated in plant cell wall degradation and interestingly, of these two α-galactosidases (PST79_3645, PST79_3417) and three β-mannosidas (PST79_2043, PST79_8128, PST79_4682) were up-regulated in germinated spores. These could potentially play roles either in fungal cell wall remodelling or degradation of plant callose. The remaining GHs were similar to enzymes required for glucan synthesis and fungal cell wall remodelling (almost all up-regulated in germinated spores), posttranscriptional modification of cell wall proteins, chitinases (almost all up-regulated in haustoria) and glycogen breakdown.

Related to cell wall modification, we also found genes with strong similarity to chitin synthases (mostly up-regulated in germinated spores) and carbohydrate deacetylases (including chitin deacetylases), which were massively up-regulated in germinated spores. Chitin deacetylases are of particular interest because they convert chitin into the less rigid chitosan, potentially avoiding recognition by plant chitin receptors [55]. The genomes of Pgt and Mlp [23] as well as the secretome of the symbiont Laccaria bicolor [56], were also found to be enriched in chitin deacetylase genes. Expression profiling of the biotrophic stage of the coffee rust Hemileia vastatrix suggested that two chitin deacetylases were most strongly expressed in the early stages of host invasion, coinciding with spore germination and tissue invasion [57]. Likewise, the rust fungus Uromyces ciceris-falae exhibits massive chitin deacetylase activity when the fungus starts to penetrate through the stomata [58]. An overall view of the data suggests complex regulation of GHs during fungal development to achieve a balance between the degradation of plant cell wall polysaccharides without triggering an immune response, and

Figure 4. Heatmaps representing relative levels of gene expression classified by metabolic categories. Relevant categories were selected from the B2G analysis. Genes belonging to these categories were listed and organized according to transcript expression values (RPKMs) in haustoria and spores. The colours were based on haustorial expression relative to the overall expression in both tissues; red indicates high expression in haustoria and green, expression in germinated spores. Specific subcategories are shown by the black arrows indicating the presence of specific transcripts discussed in the text.

doi:10.1371/journal.pone.0067150.g004
remodeling of the fungal cell wall via both degradation and synthesis to allow growth and development.

**Transport proteins.** Nutrient acquisition is presumed to be a major role of haustoria in rust infection. We therefore searched for potential nutrient transporters by BLASTx of the *Pst* contigs set against the NCBI non-redundant protein and the Transport Classification databases (TCDB: http://www.tcdb.org). The initial cut off for evaluation was e-value <1×10^-25, and the hits were further annotated manually to remove alignments of less than 100 amino acid residues. The major transporter classes that we found constituted those for amino acids and oligopeptides, sugars, small molecules and ions, and the vitamin co-factor nicotinic acid. Some of these were expressed differentially in a manner that gives insight to the pathogenic strategy of the fungus (Table S4, Transporters).

**Amino acid and oligopeptide transporters.** Pathogenic fungi have a large requirement for nitrogenous compounds for macromolecule biosynthesis. The best characterized haustorial amino acid transporters are from *U. fabae* AAT1p [59], AAT2p [22] and AAT3p [60]. In our data we found very close homologs to all three transporters (*PST79_2586, PST79_1706, PST79_2986*), each of which were expressed at very low levels in germinated spores, but expressed highly in haustoria, especially the AAT2 homolog. We also found three other putative amino acid transporters, one highly similar to AAT3 (*PST79_3074*) that was expressed strongly in germinated spores but at low levels in haustoria. Similarly to *U. fabae* [16], our data suggest that most amino acid uptake occurs via haustoria, but the high expression of one amino acid transporter in germinated spores might suggest that uptake can also occur from other cell types such as infective hyphae. It would be very interesting to test the specificity and affinity of *Pst* putative amino acid transporters; presumably the early expressed transporter should show very high affinity because of the scarcity of nutrients in the apoplast. In addition to this, eight predicted oligopeptide transporters (OPT) were found in our data, three up-regulated in haustoria (*PST79_9468, PST79_1656, PST79_8271*), one in germinated spores (*PST79_2988*) and the other four did not show expression biases. The amplification of the OPT gene family in *Pgt* and *Mlp* was proposed to be a genomic adaptation of these pathogens to obtain amino acids, nitrogen and carbon from oligopeptides derived from their host [23]. Overall our data suggest that haustoria are more active than germinated spores in uptake of amino acid compounds, although the early expression of some of those transporters could be crucial for the pathogen’s development.

**ABC and MFS transporters.** ABC transporters constitute a large superfamily of primary active transport systems that are present in all kingdoms of life, and play diverse physiological roles in trafficking of a wide range of substrates across internal and external membranes. These transporters represent one of the largest and most ancient transporter classes and derive their name from the shared highly conserved domain, the ATP binding cassette (ABC), which binds and hydrolyzes ATP [61]. MFS transporters are an ancient class of single-polypeptide secondary carriers which transport small solutes in response to chemiosmotic ion gradients [62]. We identified twelve transcripts encoding proteins with similarity to ABC transporters in the *Pst* transcriptome, but only three of these showed differential expression (*PST79_11360, PST79_10409, PST79_132*, up-regulated in germinated spores), so no generalizations about biological function were possible. In other biotrophic pathogens such as *Blumeria graminis* f.sp. * Hordei*, a marked reduction in the number of genome-encoded ABC transporters (only 20) has been correlated with the loss of secondary metabolic enzymes, reported as the lowest number known in fungi [24]. In contrast, ABC transporters seem to play an important role during pathogenesis for the hemibiotroph *M. oryzae*, especially in the appressorial stage where far more ABC transporters are expressed than in any of the studied stages of *Pst* [63,64,65]. This may reflect a need for the pathogen to exclude small molecule defense compounds secreted by the host, and to deploy secondary metabolites during tissue colonization [65]. We also identified twenty six transcripts annotated as MFS-like transporters, 16 with putative substrates, and ten classified as general MFS transporters. Of the general MFS transporters, six were expressed preferentially in germinated spores (*PST79_7841, PST79_11441, PST79_3438, PST79_8377, PST79_3969, PST79_80760*), and two in haustoria (*PST79_96, PST79_7181*). Duplessis et al [23] reported 51 and 88 MFS transporter genes in the *Pst* and *Mlp* genomes respectively, however while the total number of such transporters in the *Pst* genome is presumably similar, only a few of them were expressed during the developmental stages and sampling conditions studied here. Nevertheless, most of the MFS transporters in *Pst* were preferentially expressed in germinated spores, which is the first infectious stage that encounters host defense. This could suggest that although rust pathogens are also equipped to export toxins or metabolites, the most important roles of these proteins is in growth and development.

**Sugar transporters.** Three putative glucose transporters (MFS), seven unspecified sugar transporters (MFS), one mannose transporter and one oligosaccharide translocation protein were identified in our set of transcripts. One of the transcripts classified as a glucose transporter in *Pst* (*PST79_113*) shares 92% similarity at the amino acid level with the hexose transporter HXT1p of *U.fabae* [14]. This gene showed the highest haustorial expression of all of the sugar transporters identified, and similarly to *Uf*-HXT1, showed almost no expression in germinated spores. The expression pattern of *Uf*-HXT1p is very similar to the amino acid transporter *Uf*-AAT2p [14]. Interestingly, the haustorial expression values for the *Pst* homologs of these two transporters were also almost identical, suggesting that expression of these genes is coordinated across different rust species. We also identified three *Pst* homologs of the *Ht*-ATPase transporter *Uf*-PMA1 (*PST79_1028, PST79_3435, PST79_4481*) which is believed to be important for establishment of a proton gradient for coupling substrate translocation into haustoria [66]. Similarly to *Uf*-PMA1, all of the *Pst* homologs were expressed more strongly in germinated spores than haustoria [67]. Biochemical characterization of this *U. fabae* ATPase showed that despite the higher transcript levels in germinated spores than haustoria, its enzymatic activity is far greater in haustoria where it may act as a proton-substrate symport mechanism for *Uf*-HXT1p [67,68]. Thus, this mechanism may be conserved in *Pst* and potentially outlines a conserved transport mechanism for glucose that is key to the biotrophic lifestyle of rusts.

The remaining two putative glucose transporters (*PST79_2057, PST79_2930*) were up-regulated in germinated spores, suggesting that the fungus prepares for tissue invasion and is poised to exploit host carbohydrates as a nutritional source, or alternatively that these could export substrates for cell wall synthesis. *M. oryzae* expresses sugar transporters at an analogous stage; during appresoria formation but before tissue penetration, supporting our reasoning [63]. Finally, of the seven unspecified sugar transporters, three were up-regulated in haustoria and three were more strongly expressed in germinated spores. One of the genes expressed in germinated spores (*PST79_2314*) has a high similarity to STL1 (e-val 1×10^-61) with the *Saccharomyces cerevisiae* gene, a glycerol-proton symporter that is inactivated by glucose [69]. This might suggest that glycerol could be a primary carbon source, or
alternatively plays a role as an osmoregulator during spore germination, as has been shown for other fungi [70,71,72,73].

**Phosphate transporters.** Inorganic phosphate is an essential nutrient required for the synthesis of nucleic acids, phospholipids, cellular metabolites, and protein modification, reactions that require high phosphate concentrations [74]. We identified four transcripts encoding different phosphate transporters from our dataset. One of these was annotated as the phosphate translocator of the inner mitochondria membrane, Pho84p, one of the key proteins in this phosphate transport system, which is regulated transcriptionally in response to extracellular phosphate levels sensed by a phosphate-responsive signal transduction pathway [78]. The expression of this gene is greatly induced in low-phosphate medium and is essential for growth under low-phosphate conditions [75]. The presence of a homologous gene in Pst and its up-regulation in germinated spores could suggest that similar phosphate starvation sensing/signalling mechanisms exist in rusts, which could be crucial for cell cycle progression and growth as has been shown in yeasts [79]. In addition, thirteen transcripts similar to putative ion transporters with specificities for Ca2+, K+, Fe2+, Mg2+, Mn2+, K+ and Zn2+ were identified in our Pst data. Transporters for Zn2+ and Mn2+ were up-regulated in haustoria, while transporters for K+ and Ca2+ were up-regulated in germinated spores.

**Nicotinic acid transporters.** Nicotinic acid is an essential cofactor for many enzyme processes. We found four transcripts (PST79_362, PST79_7645, PST79_4348, PST79_3107) encoding proteins with similarities to the S. cerevisiae nicotinic acid transporter TNA1 (e-value $\leq 1 \times 10^{-30}$) [80] in our data, one of them notably up-regulated in haustoria. Vitamin B3 (nicotinamide and nicotinic acid) is essential to living cells because it can be converted to nicotinamide adenine dinucleotide (NAD$^+$), which can be reduced to NADH or the phosphorylated form NADPH. Generally, NADPH is a reductive intermediate for biosynthesis pathways, whereas NADH is an energy precursor and protects against oxidative damage. Thus uptake of host-derived niacin could be important in the haustorial stage which is greatly involved in energy production, biosynthesis of molecules for spore formation, and protection against oxidative stresses. Duplessis et al. [23] reported the _in planta_ expression of four _M. grisea_ genes with similarity to _S. cerevisaeae_ TNA1 permease at 96 hpi, supporting the idea that uptake of host nicotinic acid plays an important role during the advanced stages of the infection. We also searched for contigs that encode enzymes required for _de novo_ synthesis of niacin from tryptophan. The known pathway in organisms that use tryptophan as a source for NAD$^+$ (including yeasts, humans and some bacteria) involves six enzymes; we were able to identify five of these in our dataset, two of which were up-regulated in haustoria (Table S3, General metabolism). Thus it remains an open question whether Pst can synthesize niacin _de novo_, or it relies exclusively on its host for obtaining this essential cofactor.

**Energy use – oxidative phosphorylation.** Transcripts classified in this category were represented more in haustoria than in germinated spores (Figures 3 and 4). Transcripts with similarity to genes involved in primary pathways of energy production comprising glycolysis, citric acid cycle, and oxidative phosphorylation were mostly up-regulated in haustoria (Table S5, Energy metabolism). Similar observations have been made for the obligate biotrophs _U. fabae_ [81] and _Blumeria graminis_ [82] where genes related to glycolysis are up-regulated during the parasitic stage. Those few Pst genes with higher expression in germinated spores corresponded to the key enzymes of the glyoxylate cycle, which will be described later. In addition, a phosphate translocator of the inner mitochondria membrane (PST79_4938), which carries out the coupled transport of H$_2$PO$_4^-$ and H$^+$ for subsequent ATP synthesis via the electron transport chain, was massively up-regulated in haustoria, consistent with a high demand for ATP in this tissue. This could be a consequence of the different nutritional status of haustoria and germinated spores. While the stored energy resources are limited in spores, haustoria can derive sugars abundantly from the host to feed energy production pathways and drive biosynthetic pathways for subsequent development (Figure 5).

**Fatty acid metabolism.** Acetyl-CoA is the basic currency of carbon metabolism within the cell. It is the immediate product of carbohydrate catabolism and β-oxidation of fatty acids, and when in excess, can be used for biosynthetic processes through the glyoxylate cycle, gluconeogenesis and glyceroneogenesis pathways. In our transcriptomic study, the gene for triacylglycerol lipase (PST79_8009) together with nearly all those for the enzymes of glyoxylate cycle, gluconeogenesis and glyceroneogenesis were significantly up-regulated in germinated spores compared with haustoria (Table S3, General metabolism). This suggests that the activity of metabolic pathway for converting fatty acids into carbohydrates is important in the initial stages of infection. Key genes involved in β-oxidation of fatty acids such acyl-CoA dehydrogenase (PST79_1352) and enoyl-CoA hydratase (PST79_4902) were also up-regulated in germinated spores, suggesting that once triacylglycerides are broken down, the released fatty acids are oxidised and the acetyl-CoA produced can be fed into gluconeogenesis by means of the glyoxylate cycle. The two key enzymes in the glyoxylate cycle, isocitrate lyase (PST79_6444) and malate synthase (PST79_4401), were highly up-regulated in germinated spores. However, malate dehydrogenase and citrate synthase, two other enzymes in this cycle, did not show the same biased expression pattern between the two tissue types. Malate dehydrogenase participates in the glyoxylate cycle and is also involved in the malate-aspartate shuttle, which is the system responsible for translocating electrons produced in glycolysis through the mitochondrial membrane for oxidative phosphorylation. Thus, the role of this enzyme is expected to be important in germinated spores but also in haustoria because of their high glycolytic activity. Apart from fatty acids, the second product of the triacylglycerol breakdown is glycerol. This compound can be converted to intermediates of gluconeogenesis by means of glycerol kinase (PST79_6140) and glycerol-3-phosphate dehydrogenase (PST79_3414), genes which also showed significant upregulation in germinated spores. As mentioned previously, a gene similar to _STLI_ (encoding a glycerol transporter) was also up-regulated in germinated spores, which suggests that glycerol derived from the host apoplast could be taken up and incorporated via the gluconeogenic pathway. Finally, genes encoding enzymes involved in glycogen breakdown were also up-regulated in germinated spores, which also results in glycerol production. In _Pst_ uredinospores, as well as spores of other plant fungal pathogens, glycogen is a stored energy source that is utilized during germ tube extension [83,84]. These observations are consistent with data from other biotrophic and hemibiotrophic pathogens.
fungi such as *B. graminis* and *M. oryzae* [65,82], where similar metabolic activity has been detected in early stages of the infection. Overall, our results point towards spore germination utilising stored compounds including lipids and glycogen as sources for energy production (Figure 6), whereas the massive upregulation of *HXT1* and enzymes of glycolysis, TCA cycle and oxidative phosphorylation in haustoria suggest that host-derived glucose is the primary carbon source.

**Pentose Phosphate Shunt (PPS).** The PPS is an alternative fate for glucose which produces NADPH and ribose-5-phosphate. Ribose-5-phosphate can be used for nucleotide synthesis (oxidative branch), or recycled for energy production via glycolysis (non-oxidative branch). An important product of the PPS is the reductant NADPH, which is used almost exclusively in anabolic pathways, for example biosynthesis of fatty acids, amino acids and nucleotides including ATP. We found evidence for expression of all enzymes of the PPS in both germinated spores and haustorial (Table S3, General metabolism). Interestingly, all of the enzymes participating in the oxidative branch of this pathway showed an expression bias in haustoria as did transaldolase, which belongs to the non-oxidative branch. In germinated spores, the genes encoding the enzymes ribulose-phosphate 3-epimerase (PST79_10575) and phosphoglucomutase (PST79_5659) were up-regulated. The ability of this pathway to work in different directions suggests that it could play different roles in each tissue type. While both tissues are expected to have a strong requirement for biosynthetic processes, this is expected to be higher in haustoria which presumably contribute to provision and transport of substrates to fuel the growing infection and eventual spore biogenesis (Figure 5). Thus, the haustorial PPS seems to engage the oxidative arm for maximal NADP⁺ reduction and ribose-5-phosphate production. Additionally, an important enzyme of the non-oxidative branch, transaldolase, showed higher expression in haustoria than in germinated spores. The non-oxidative branch can convert five carbon sugars into six and three carbon sugars, which are fed into glycolysis. This would appear to be unnecessary in haustoria which should not be limited for glucose. Alternatively, the non-oxidative arm of PPS may work in the opposite direction, taking fructose-6-phosphate and glyceraldehyde-3-phosphate generated by glycolysis to produce ribose-5-phosphate, which is the basis of production of ATP, NAD⁺ and nucleic acids. However, the remainder of the enzymes that participate in the non-oxidative branch did not show a clear expression trend, which precludes clear conclusions on the role of this branch of the PPS. In *Plasmodium falciparum*, a human pathogen with an intracellular niche that is comparable to the rust haustorial stage, the oxidative arm is thought to operate during early stages of parasite development while the non-oxidative arm is active later in the cycle [85,86]. These studies also showed that transcription of the PPS genes is not always coordinated, with activity of the pathway as a whole often dependent on expression of a single gene whose transcription is the rate limiting step for deployment of the pathway’s activity. In *B. graminis*, the absence of clear coordination in expression of the enzymes of PPS has also been observed [82]. These discrepancies

**Figure 5. Metabolic processes in haustoria.** Metabolic processes or specific enzymatic activities overrepresented in haustoria identified by B2G analysis and manual annotation are highlighted in this cartoon. Orange boxes and yellow ovals are metabolic processes or particular enzymes that showed statistically significant upregulation in haustoria. Light orange boxes are metabolic processes where some genes showed statistically significant upregulation in haustoria, and others showed a tendency to be more expressed in haustoria but were not statistically significant. doi:10.1371/journal.pone.0067150.g005
require direct biochemical investigation which will provide more
views about the physiological implications of these patterns.

**Thiamine biosynthesis.** A number of metabolic enzymes
require the vitamin B1 derivative thiamine pyrophosphate as a
cofactor. Examples from the pathways described here include
alpha-ketoglutarate dehydrogenase (α-KGDH) in the TCA cycle;
transketolase in the non-oxidative PPS; and pyruvate dehydroge-
nase (PDH) which connects glycolysis and the TCA cycle. We
found that transcripts associated with the thiamine biosynthesis
pathway were massively up-regulated in haustoria, with the
exception of the genes for thiamine-phosphate kinase and
thiamine pyrophosphokinase which could not be found in our
data. These enzymes are responsible for phosphorylating thiamine
phosphate and thiamine diphosphate respectively, and interest-
ingly, the latter is also absent in the genome of the malarial
parasite *P. falciparum* [85]. The genes we found include those
encoding homologs of thiamine biosynthesis genes *THH1* (4-amino-
5-hydroxymethyl-2-methylpyrimidine phosphate synthase)(PST79_1423, PST79_612, PST79_1422) and *TH2* (hydrox-
yethylthiazole phosphate synthase)(PST79_882) from *U. faber*
identified previously by Hahn and Mendgen [13]. The massive expression of these genes almost exclusively in haustoria agrees with studies done in other rusts [23,28], but it is not clear why the pathogen requires such high expression at this stage, which seems excessive if the role is simply to produce sufficient cofactors for metabolic enzymes. One explanation may be that thiamine is exported to other tissues. Alternatively, thiamine is known to alleviate stress in different organisms as an antioxidant. It is interesting that in the oomycete, all haustorium-forming species have lost the thiamine biosynthetic pathway [87], which suggests that they must obtain thiamine from the host.

**Nitrogen metabolism.** Previous genome studies of obligately biotrophic fungal plant pathogens have noted the absence of genes encoding enzymes for nitrate and sulfur assimilation [23,24]. Likewise, we were unable to identify genes for nitrate reductase or nitrite reductase, or for a nitrate transporter that could import nitrate directly from host tissue. Instead, we found a transcript encoding a protein with very high similarity to an ammonium transporter (PST79_5294) expressed in both germinated spores and haustoria. Ammonium is the preferred form of nitrogen for most organisms [88] and is assimilated into glutamate and glutamine. Glutamate synthase and glutamine synthetase are the key enzymes involved in ammonia assimilation. Glutamate is formed from α-ketoglutarate and glutamine in a reaction that is catalyzed by glutamate synthetase, followed by amination of glutamate by NH$_4^+$ to form glutamine catalyzed by glutamine synthetase. Subsequently, glutamate is used to synthesise the majority of amino acids, whereas glutamine serves as an amino-donor during other biosynthetic processes. Interestingly, the transcripts encoding genes for both glutamine (PST79_3032) and glutamate synthase (PST79_2575) identified here were up-regulated in haustoria, consistent with a greater role for biosynthesis in this organ. Genes encoding other enzymes that participate in subsequent steps including glutamate dehydrogenase (PST79_536), aspartate aminotransferase (PST79_7929), asparagine synthase (PST79_1116) and asparaginase (PST79_83), were all expressed in both tissues, suggesting nitrogen compounds obtained from the host can be converted freely into compounds required for pathogen growth and development.

**Sulfur metabolism.** Sulfur is an essential component of living cells as it is a fundamental component of the amino acids methionine and cysteine, Coenzyme A, and iron-sulfur enzymes. Most fungi take up sulfur as sulfate which is reduced to sulfide as a precursor of cysteine [89]. Despite this, the steps for sulfate assimilation are energetically expensive and fungi prefer to utilize cysteine or methionine. To date, enzymes for sulfate uptake and reduction have not been identified in obligate biotrophs such as B. graminis [90], P. graminis f. sp. tritici [23]. In addition, genes encoding enzymes for sulfate reduction including sulfite reductase and phosphoadenosine phosphosulfate reductase were absent, consistent with observations for other obligate plant pathogens. However, we found a contig (PST79_3132) that encodes a transporter with very high similarity to S-methylmethionine (SMM) permease, which was almost exclusively expressed in haustoria compared to germinated spores. This is interesting because SMM is present at very high levels in wheat phloem [91] and as such, could be an abundant source of sulfur metabolites for the parasite. Apart from this, we were unable to identify candidate cysteine or methionine transporters in our data. Lastly, we identified a gene encoding cysteine synthase (PST79_10177), which assimilates reduced sulfur (S$^{2-}$) into cysteine, expressed at low levels in both tissues. Genes with high similarities with those involved in the interconversion of homo-cysteine and cysteine through the intermediary formation of cystathionine (transsulfuration pathways) were also identified. This pathway creates cysteine from methionine.

**Glutathione metabolism.** Glutathione provides redox buffering to cells, and plays a role in many cellular processes including iron metabolism [92]. The anti-oxidant properties of glutathione contribute to anti-defence against host active oxygen compounds. Glutathione is derived from the amino acids glycine, glutamate and cysteine. The two key enzymes for synthesis of glutathione are glutamate-cysteine ligase (PST79_3565) and glutathione synthetase (PST79_1176). Both of these genes showed greater expression in haustoria, and PST79_3565 was significantly upregulated. This could suggest an ongoing need for antioxidants in the static haustorial stage which is exposed to plant defences. We also note that glutathione reductase (PST79_10658) and glutathione peroxidase (PST79_5949)enzymes involved in the redox balance) as well as genes involved in the response to oxidative stress were more highly expressed in germinated spores than in haustoria (Table S3, General metabolism). All this would suggest the presence of glutathione in spores prior to development. Thus, one function of haustoria could be to synthesise glutathione for subsequent deposition into new spores as a pre-formed environmental defence.

**Cytochrome P450s.** Cytochrome P450s comprise a diverse superfamily of proteins containing a heme cofactor that catalyse the oxidation of organic substances. Among their most common substrates are steroidal molecules as well as drugs and toxins. The P450 system is important in fungal evolution for adaptation to ecological niches [93], and its potential role in compound detoxification including metabolism of plant defense molecules [13,94]. A total of 29 and 17 P450s are encoded by the genomes of M. larici-populina and P. graminis respectively, but the current dataset, we identified only eight expressed P450 genes, five of which were up-regulated in haustoria (Table S3, General metabolism). One of the P450s (PST79_8680) we identified is highly similar to CYP51A1, a lanosterol 14 alpha-demethylase, required for conversion of lanosterol to cholesterol. This enzyme is targeted by the commonly used azole class of antifungal drugs, which occupy the active site of the enzyme inhibiting the production of ergosterol, a component of the fungal cell membrane [95]. Two other P450s (PST79_1792, PST79_1793) up-regulated in haustoria belong to the CYP67 family, which was first reported in U. fabae as an in planta-induced gene [13]. The remaining genes encode unspecified mono-oxygenases.

**Transcription factors.** Overall, our analysis revealed approximately 37 transcripts encoding proteins related to transcription factor activity, from both haustoria and germinated spores (Figure 4 and Table S6, HeatMap). Despite this, it is difficult to obtain specific insight into the patterns and functions of these genes. However, there were a few exceptions where we could predict a possible function based on BLAST2GO and BLASTx searches (Table S1, General metabolism). One of the most interesting genes found in our data set was a Ste12-like gene (PST79_9215), which was up-regulated in haustoria. Ste12 was first identified in a yeast sterile mutant, and multiple studies have shown that Ste12-like genes play major roles in regulating morphogenetic programs in response to environmental changes (see review [96]). Yeast Ste12 is an important regulator of invasive growth and pseudohyphal development [97] and its homologs in a number of fungi are important for sexual development and pathogenicity [98,99,100,101]. In the root pathogen Fusarium oxysporum, a ste12 mutant was impaired in pathogenesis on tomato and the pathogen could not differentiate into specialized infectious structures [102]. In the hemibiotrophic fungal pathogens Colletotrichum lagenarium, Colletotrichum lindenuthianum and M. oryzae, Ste12-like genes are important for penetration of leaf surfaces from
upregulation of this gene in haustoria could suggest these sources limited [110]. The presence of a similar gene in catabolic enzymes for utilization of secondary sulfur sources when plants, the mutant strain acquired resistance to three groups of [118]. Although disruption of component histidine kinase originally isolated from germinated spores (PST79_186) has high similarity to Os-1, a two-component signalling systems participate in processes such as diverse protein kinases were most highly represented. In fungi, protein superfamily of guanine nucleotide exchange factors, and environmental sensing, oxidative stress response, cell-cycle control, two-component signalling systems and transcription (PST79_5377), an MCM-domain-containing protein specifying haustorial development.

Lastly, five other genes associated with transcription factor activity were found in our data, encoding a putative CGAAT-box binding factor (PST79_5498), a putative activator of basal transcription (PST79_5377), an MCM-domain-containing protein (minichromosome maintenance ATPase) (PST79_3174), a putative bZIP transcription factor (PST79_8536) and a gene with similarity to the regulator Cys3 (PST79_492) reported originally in Neurospora crassa [109]. Cys3 controls the synthesis of a set of catabolic enzymes for utilization of secondary sulfur sources when sulfur-containing amino acids and inorganic sulfate are missing or limited [110]. The presence of a similar gene in Ps might suggest that secondary sulfur sources play a role in biotherapy, and upregulation of this gene in haustoria could suggest these sources are derived from the host.

Signal transduction. Cellular differentiation and filamentous growth are finely regulated in many fungi in response to environmental and nutritional signals. External stimuli must be transmitted within the fungus by signal transduction pathways to enable the appropriate response. Our transcriptomic analysis revealed high expression of signal transduction components in both germinated spores and haustoria. Transcripts encoding proteins related to two component response regulators, the Ras protein superfamily of guanine nucleotide exchange factors, and diverse protein kinases were most highly represented. In fungi, two-component signalling systems participate in processes such as environmental sensing, oxidative stress response, cell-cycle control, and switching between non-pathogenic and pathogenic states [111,112,113]. They are typically comprised of a sensor-kinase protein, a phosphorylase protein, and a response regulator. We found genes encoding two-component regulatory proteins in both haustoria and germinated spores (Table S6, HeatMap). Generally, the proteins expressed in each tissue were distinct, consistent with the idea that they respond to different stimuli. There was a larger representation of these genes in germinated spores suggesting a more complex interpretation of environmental signals prior to development of the haustorial structure. One gene up-regulated in germinated spores (PST79_186) has high similarity to Os-1, a two-component histidine kinase originally isolated from Neurospora crassa. Os-1, also known as Nik-1, is required for adaptation to high osmolality [114,115]. Three groups of selective fungicides target the osmotic stress signal transduction pathway [116,117]. Intracellular osmotic pressure plays important roles in fungal development, for example as inside the M. oryzae appressorium [118]. Although disruption of HIK1 (Os-1) in this pathogen caused no defect in growth on normal media or in pathogenicity to rice plants, the mutant strain acquired resistance to three groups of fungicides (phenylpyrroles, dicarboximides, and aromatic hydrocarbons) [119]. The almost exclusive expression of this gene in Ps germinated spores suggests that the osmosensitive signal transduction pathway could play a role in the early stages of stripe rust disease.

Ras proteins are a superfamily of guanine nucleotide exchange factors in which the phosphorylation status of the bound nucleotide operates a binary switch. They have numerous signalling roles in cellular functions including cytoskeletal integrity, proliferation, differentiation, cell adhesion, and cell migration. Ras proteins are active in the GTP-bound form, which is promoted by guanine nucleotide exchange factors (GEF) in an exchange reaction with GDP, and deactivated by GTPase activating proteins (GAP) that promote dephosphorylation of the bound GTP nucleotide. We identified evidence of expression of genes encoding proteins with similarity to Ras proteins called ADP ribosylation factors (ARF) in both sampled tissues, which is of particular interest to rust pathogenesis because they regulate vesicular traffic, phospholipid metabolism and actin remodeling. Other genes encoding other types of Ras protein as well as GAPs and GTPases were also found in both tissues, suggesting complete representation of these complex molecular switches that regulate cell fate.

Conclusions

Transcriptome studies such as the current work considerably enhances analysis of genome information. The 454 platform used here gives significantly longer reads than Illumina, which both helps with transcript assembly and demonstrates gene expression. In contrast, small-read based assemblies are subject to assembly errors and the uncertainties of gene prediction, although these also benefit from annotation with transcriptomic data. One important aspect of defining transcriptomes, as described here, is that it cuts through the complexity of the genome. Thus, while the predicted coding capacity of rust genomes is ~22,000 genes, we identified ~12,000 expressed contigs in the sampled tissues. Presumably, many of the remaining genes are expressed in different contexts (for example on the alternate host, or as part of different spore stages), or were simply undetected or unassmebled in our study. Although isolation of haustoria is a very powerful way to track the pathogenic stage of the fungus, we did not analyse infected tissue which includes the infectious hyphal stage that presumably also secretes effector proteins. Another important aspect of our study is quantification of expression using Illumina data, which considerably enhances the accuracy and reliability of our expression measurements. Thus, use of 454 and Illumina together provides a robust expression data set. This allowed us to predict a suite of 497 potential effector genes, defined chiefly as genes upregulated in haustoria that encode secreted proteins. We also note that not all secreted proteins will be virulence effectors, and some within our dataset show clear hallmarks of roles in fungal cell wall modification.

The current data provide a comprehensive view of the stripe rust transcriptome in germinated spores and haustorial stages. Overall, the fungus seems to use similar biochemical pathways to those described in better characterised species such as M. oryzae, but with adaptation to the particular pathogenic lifecycle of wheat rusts. Haustoria and spores are clearly working in different ways. The germinating spores polarises growth from stored energy until such time as it is able to mine its own resources, whereas haustoria are a sessile stage devoted to nutrient extraction and defence suppression. The urediniospore contains vast stores of energy in the form of lipid bodies, sugar alcohols and proteins [120,121] to allow the foraging germling to find the pathogenic
niche, and in that sense is analogous to a plant seed. From our data, we can see that haustoria create abundant demand for sugars, amino acids, and nitrogen, through high expression of transporters for these fundamental building blocks. In addition, we propose a new mechanism for sulphur assimilation by uptake of methylmethionine. Sugars are used immediately by haustoria for the production of ATP through glycolysis, TCA cycle and oxidative phosphorylation. By contrast, spores seem to obtain energy by utilising lipid resources via the glyoxylate pathway. A further important haustorial function is biosynthesis of macromolecules, and it is clear that steady production of the redundant NADPH is made possible by import of nicotinic acid through dedicated transporters, and its ultimate conversion to NADPH by the PPS. Spores also use the PPS but it is likely that this is works in the opposite (oxidative) direction, for energy production. Molecules made in haustoria are presumably exported directly to spores for biogenesis and storage, so the molecular pathways active in haustoria should reflect spore composition directly. With respect to the completeness of our dataset, we note as have previous authors that many genes are novel so cannot be annotated. For this reason, it is difficult to exclude the presence of certain genes with major implications for the biotrophic lifecycle, such as N and S assimilation. Of course, the absence of these genes could also be ascribed to low expression of miss-assembly of small reads. These are important challenges for the field because barring technological breakthroughs, most fungal genome sequences are likely to have large components of small sequence reads in their genesis.

Materials and Methods

Strains and Culture Conditions

Seedlings of the Pst-susceptible wheat cultivar Morocco were grown in the greenhouse under 70% relative humidity at 21°C and 16:8 light:dark cycle. Seven-day old seedlings were inoculated with freshuredidospores of P. striiformis Lsp. tritici strain 104E137A-1, and incubated for 48 h in 100% humidity at 9°C in the dark. Subsequently plants were transferred to a growth chamber at 17°C with a 16:8 light cycle. About 20 grams of heavily infected tissue were collected 9 days after infection (dai; just prior to sporulation). Germinated spores were collected 9 days after infection (dai; just prior to sporulation) and immediately processed for haustoria isolation. Germinated spores were obtained by germination of about 80 mg of fresh spores harvested from infected leaves 17 dai on sterile distilled water at 9°C for 15 h in the dark. Germinated spores were collected by filtration with an 11 μm nylon mesh, and frozen in liquid nitrogen and stored at −80°C prior to extraction of total RNA.

Haustoria Isolation

For 454-based transcriptome analysis, haustoria were isolated from stripe rust-infected Morocco leaves at 9 dai by Concanavalin A affinity chromatography (Text S1 and Figure S4) [122]. Fifteen separate preparations were made each from 20 g of infected wheat leaves that were sequentially washed with chilled tap water, 2% bleach, water, 70% ethanol and Milli-Q purified water. Haustorial isolation was performed as described [123] using a final 20 μm pore nylon mesh to remove the bulk of the plant cell material before affinity purification. Purified haustoria were pelleted by centrifugation at 14,000 g for 5 min, frozen in liquid nitrogen and stored at −80°C prior to RNA isolation.

For Illumina sequencing, approximately 10 μg of total RNA per biological replicate of isolated haustoria and germinated spores were processed with the mRNA-Seq Sample Preparation kit from Illumina to produce the sequencing libraries. Quality and quantity controls were done on Agilent 2100 Bioanalyzer using a DNA 1000 chip kit and each library was diluted and used for sequencing using an Illumina Genome Analyser GX II platform. Libraries from haustoria samples and germinated spores samples were sequenced with 100-base paired-end reads.

Whole Transcriptome Analysis by 454 and Illumina Sequencing

CLC Genomic Workbench 4.0 software (http://www.clcbio.com/) was used for de novo assembly of 454 reads, prediction of ORFs, Illumina read mapping against reference contigs, computation of normalized counts (expressed in reads per kilobase of exon model per Million mapped reads – RPKM [41]) and differential expression analysis. The parameters used for de novo assembly were similarity 0.97; length fraction 0.5; insertion cost 3; deletion cost 3; mismatch cost 2. The mapping of Illumina reads
against reference contigs was done after removing adapter sequences and reads of low quality (trim quality score 0.05, max nucleotide ambiguities 2 and minimum number of nucleotides in reads = 35), allowing up to two base mismatches. Reads that mapped to multiple sites were assigned randomly and proportionally to one of the mapped sites. The random distribution was done proportionally to the number of unique matches that the genes to which it matches have, normalized by the transcript length.

Three biological replicates of each tissue were used for differential gene expression analysis. A 0.84–0.88 Pearson correlation coefficient was obtained for haustorial biological replicates and 0.88–0.93 for germinated spores biological replicates. The differential expression between haustoria and germinated spores was evaluated using Baggerley’s test [42] by treating the same types of sample as one group. The genes with a false discovery rate (FDR) corrected $P$-value less than 0.05, a fold change $>2$ and a difference of at least 20 were considered to be significant. Short reads for the 454 and Illumina datasets were deposited in the NCBI Sequence Read Archive under the accession numbers SRR579533-40.

For the gene ontology (GO) classification the reference transcripts set created by pooling raw reads from both transcriptomes was analyzed using BLAST2GO 2.5.1 [36]. Parameters were set to maximum $e$-value $<10^{-23}$, maximum number of alignments to report $=20$ and highest scoring pair length $=33$ amino acids. BLAST2GO was also used for GO functional enrichment analysis of the transcripts differentially expressed in both germinated spores and haustoria, by performing Fisher’s exact test with false discovery rate (FDR) correction to obtain an adjusted $P$-value. Additionally, BLASTn and BLASTx screening against NCBI non-redundant nucleotide and protein databases were used for manual analysis. Only the transcripts that belonged to one of the functional categories were analysed manually to verify if more than one transcript represented the same gene, in which case it was annotated in Tables S3–S6. For the remainder of the transcripts, which comprise anonymous sequences, this was not possible because different transcripts might have the same top hit without meaning that they were derived from the same gene. The haustorial transcriptome was further analysed with SignalP 3.0 [33] for signal peptide prediction, TMHMM 2.0 [34] for discarding predicted secreted proteins with transmembrane domains and TargetP [35] for discarding predicted secreted proteins with predicted mitochondrial location.

RT-PCR

Total RNA was isolated from infected (9 dai) and uninfected wheat leaves ground in liquid nitrogen and extracted using QIAGEN (Doncaster Australia) Plant RNeasy kit according to the manufacturer’s instructions. For cDNA preparation, 2 μg of total RNA were mixed with 1 μl of oligo(dT)$_{18}$ and Milli-Q water up to 11 μl and heated at 70°C for 10 min, followed by cooling on ice for one minute. DTT was added to a final concentration of 1 mM together with 1 μl of dNTPs (10 mM), 4 μl of 5× superscript II buffer (supplied with superscript) and 0.5 μl of SuperScript III reverse transcriptase (200 U/μl, Invitrogen). After incubation for 1 h at 42°C, the reaction was stopped by heating 15 min at 70°C and 1 μl of RNase H (1 U/μl) was added to each tube and incubated for 20 min at 37°C. Target cDNAs were amplified using a dilution of 1:15 of the synthesized cDNA as template, specific primers forward and reverse and 2 μl PCR Master Mix (Promega). PCR was performed using a MiniCycler (MJ Research) and consisted of 30 cycles of denaturation at 94°C for 30 sec, annealing at 58°C for 30 sec and extension at 72°C for 1 min.

Supporting Information

Figure S1 Number of 454 reads per assembled contig for germinated spore and haustorial samples. (TIF)

Figure S2 Length of contigs for germinated spores and haustorial samples assembled from 454 data. (TIF)

Figure S3 Coverage of 454 contigs assembled from 454 data from germinated spores and haustorial samples. (TIF)

Figure S4 Isolated haustoria by affinity chromatography. A. Stripe rust haustoria isolated by affinity chromatography. Haustoria isolated by ConA affinity chromatography using a 20 μm mesh to filter homogenized tissue and remove plant cell debris. Bright field image with differential-interference contrast, haustoria (black arrows) and contaminating chloroplasts can be seen. B. Fluorescence microscopy of isolated haustoria by affinity chromatography. Haustoria were isolated from infected tissue by ConA affinity chromatography as described in Materials and Methods. Bright field and fluorescent images showing isolated haustoria after 30 min incubation with WGA-FITC or ConA-Alexa 594. All images were collected on a Leica DMR epifluorescence microscope. (TIF)

Table S1 Transcripts reference set. 12,282 EST contigs assembled from germinated spores and haustorial datasets were used as references to map Illumina RNA-seq data from the same tissues. Three biological replicates for haustoria (H1, H2, H3) and spores (S1, S2, S3) were used for sequencing. Differences between each tissue were established by RPKM values for each replicate which were subjected to Baggerley’s test to provide statistical support for differences in gene expression. (XLSX)

Table S2 Haustorial secreted proteins (HSPs). A compilation of analyses of the 437 predicted haustorial secreted proteins. The table shows values for expression levels based on Illumina data (RPKM values, green for haustoria and orange for germinated spores), differential expression, cysteine content, BLASTx analysis against the NCBI-nr protein database, BLAST2GO analysis, PFAM motif analysis and Y/F/WxC motif content. In the Y/F/WxC column, 1b30 indicates a single of these motifs before the 30th amino acid, whereas 1a30 represents a single motif after the 30th amino acid. (XLSX)

Table S3 General metabolism. Transcripts encoding proteins with similarities to enzymes involved in metabolic pathways, or with particular roles as described in the text. Expression levels based on Illumina data are presented (RPKM values, green for haustoria and orange for germinated spores) as well as BLASTx analysis against the NCBI-nr protein database. (XLSX)

Table S4 Transporters. Transcripts encoding proteins with similarities to different transporter categories were identified by BLASTx against the NCBI-nr protein database and the Transport Classification Database (TCDB). Expression levels based on Illumina data are presented (RPKM values, green for haustoria and orange for germinated spores). (XLSX)

Table S5 Energy metabolism. Transcripts encoding proteins with similarities to the enzymes of glycolysis, citric acid cycle, and
oxidative phosphorylation, and their in silico expression levels based on Illumina data are presented (RPKM values, green for haustoria and orange for germinated spores). (XLSX)

Table S6 Gene expression values for the ontological analysis presented in Figure 4. Transcripts analysed with BLAST2GO and classed into different metabolic categories summarised in Figure 4 are listed in this table. The BLASTx values against the NCBI-nr protein database are included and their in silico expression levels based on Illumina data are presented (RPKM values, green for haustoria and orange for germinated spores). Transcripts highlighted in brown were excluded from the analysis due to low RPKM values (<10). (XLSX)

Table S7 Illumina sequencing data before and after mapping against the transcripts reference set. Three biological replicates were sequenced with Illumina for isolated haustoria (H) and germinated spores (S). The table shows the millions of reads obtained per replicate and the percentage of reads mapping against the transcripts reference set assembled from 454 haustoria and germinated spore data. (DOCX)

Table S8 List of primers to validate candidate effector gene expression by non-quantitative RT-PCR.

References


120. Staples RC, Wynn WK (1965) The Physiology of Uredospores of the Rust Fungi. Wheat Stripe Rust Transcriptome during Infection

Wheat Stripe Rust Transcriptome during Infection
Chapter 8

Purification of Fungal Haustoria from Infected Plant Tissue by Flow Cytometry

Diana P. Garnica and John P. Rathjen

Abstract

A hallmark of biotrophy in many fungal plant pathogens is the formation of a specialized pathogenic structure called the haustorium from infectious hyphae. This is the major parasitic structure, where nutrients are taken up from the host and pathogenicity factors are exported to the host tissue. Obligate biotrophic fungi can typically be cultured in vivo only to a limited extent and do not produce haustoria under these conditions. This has hampered the application of classic molecular biology techniques to haustoria-forming pathogens. The lectin Concanavalin A (Con A), which binds specifically to sugars present on the exterior of rust haustoria, was first used in a column-based affinity purification procedure in 1992 (Hahn and Mendgen, Protoplasma 170:95–103, 1992). Here we describe a new technique where we combine initial gradient purification of haustoria with flow-sorting based on labeling of haustoria with fluorescent Con A. Our method allows haustorial isolation with purity above 98% and yields ten times more isolated haustoria in a single experiment than the previous procedure.

Key words Rust fungi, Puccinia, Haustoria, Density gradient, Flow cytometry

1 Introduction

Obligate fungal parasites source nutrients from living cells of their plant hosts through a sophisticated cellular structure termed the haustorium [1]. After penetration of host tissue by the fungus, the invasive hyphae contact host cells and start to differentiate. The fungus perforates the host cell wall and expands within, invaginating the host plasma membrane but not breaching it, to form a single-lobed or multilobed structure. The haustorium remains separate from the host cell by the extrahaustorial membrane (EHM), which is derived from and contiguous with the host plasma membrane, but appears to be differentiated from it [2]. On the fungal side of the EHM is a gel-like structure called the extrahaustorial matrix, composed largely of polysaccharides, which is bordered by the fungal plasma membrane on its interior surface. The haustorium remains connected to the fungal body through a
neck structure which contacts the EHM, thus sealing the extra-
haustorial matrix from the plant cell. Molecular exchange between
the fungus and plant cells must occur across the extrahaustorial
matrix; however, the structure and function of this interface is
poorly understood [2].

Haustoria are not only feeding structures; they induce struc-
tural changes in the host cell including cytoskeletal rearrange-
ments, nuclear migration, and chromatin condensation [3], and
there is evidence that they influence host cell metabolism [4, 5].
Furthermore, they deliver essential virulence molecules called
“effectors” into the extrahaustorial matrix, several of which are
subsequently translocated into host cells [6, 7]. As such, the haus-
torium is a site of concerted host-pathogen interaction, and
describing its functions is essential to understanding biotrophy.
Despite its importance, its inaccessibility and the inability to cul-
ture it in vitro has constrained experimentation. Hahn and
Mendgen [8] found that lectins showed differential affinity for the
haustoria of *Uromyces* spp. and *Puccinia* spp., with Concanavalin A
(Con A) showing the highest affinity. Additional observations led
to the conclusion that Con A recognizes α-linked mannoside resi-
dues on the haustorial wall and the extrahaustorial matrix, since the
EHM appears to be lost during isolation of haustoria from leaf tis-
sue. Affinity of Con A for surface sugars is the basis of a scheme to
purify haustoria by column-based chromatography from crude tis-
sue extracts [8]. However, this purification method has two major
disadvantages: the high level of contamination by chloroplasts
which far outnumber haustoria, and the low yield of haustoria.
Moreover, although the method has been used for other rust fungi
[9], it failed to purify the haustoria of the powdery mildew species
*Sphaeroteca fuliginea*, *Erysiphe pisi*, and *E. graminis* f.sp. *hordei*,
suggesting that haustoria of these species differ in their external
composition. Alternative separation techniques such as density gra-
dient centrifugation have been used to isolate haustoria of pow-
dery mildew fungi [2, 10, 11]. However, none of these techniques
achieve high levels of purity, which is necessary for subsequent
applications such as transcriptome sequencing, proteomics analy-
ysis, and metabolomic studies.

Recently, Takahara et al. [12] described a method for fluores-
cent vital staining of the intracellular hyphae of the pathogen
*Colletotrichum higginsianum* from homogenates of infected
Arabidopsis leaves, which could then be purified by fluorescence-
activated cell sorting (FACS). Here we describe the purification of
wheat stripe rust haustoria by Percoll density gradients combined
with FACS based on affinity staining of the haustoria with fluores-
cent Con A. The method generates high yields of essentially pure
haustoria that are suitable for downstream analyses that demand
high purity.
2 Materials

Prepare all solutions using sterile MilliQ water and analytical grade reagents.

2.1 Haustoria Isolation from Plant Tissue

1. Infected plant material (see Note 1).
2. Ethanol 70%. Prepare 250 ml and store at 4 °C.
3. Two liters of chilled MilliQ water.
4. Sodium hypochlorite 2% (v/v), prepare fresh.
5. Sterile paper towel.
7. Homogenization buffer: 0.2 M sucrose, 20 mM MOPS pH 7.2, and 0.2% (v/v) β mercaptoethanol, added freshly. Prepare 250 ml of this buffer and store at 4 °C.
8. Isolation buffer (1× IB): 0.2 M sucrose, 20 mM MOPS pH 7.2. Prepare 500 ml of this buffer and store at 4 °C.
9. Isolation buffer (10× IB): 2 M sucrose, 0.2 M MOPS pH 7.2. Prepare 100 ml of this buffer and store at 4 °C.
10. Two clean and sterile 1 L Erlenmeyer flasks.
11. Ten Oak Ridge round-bottom centrifuge tubes, 45 ml capacity, with polypropylene screw closure.
12. Two nylon meshes (~20 cm × 30 cm), one of pore size 100 μm and the other of 20 μm (see Note 2).
13. Four polycarbonate round-bottom tubes, 45 ml capacity, non-lipped.
15. Refrigerated benchtop centrifuge with swingout rotor.
16. Concanavalin A, Alexa Fluor® 488 Conjugate (Molecular Probes® C11252, see Note 3).
17. Rotary mixer.
18. Glass vial with plastic lid, 6 ml capacity.

2.2 Purification of Haustoria by Flow Cytometry

1. Dickinson BD FACSARIA II cell sorter (BD Biosciences), or similar.
2. 5 ml glass assay tubes.
3. Phosphate buffered saline (1× PBS): 8 g NaCl, 0.2 g KCl, 1.44 g Na₂HPO₄, 0.24 g KH₂PO₄, dissolved in 1 L of MilliQ water, final pH 7.4, and sterilized by autoclaving.
3 Methods

Carry out all procedures at 4 °C unless otherwise specified.

3.1 Haustoria Isolation from Infected Tissue Using Percoll Gradients

Steps including the timing of tissue harvesting (see Note 1), the amount of infected tissue to be processed, the pore size of the meshes, and Con A staining time, have to be determined empirically according to the pathosystem under investigation.

1. Harvest 20–25 g of heavily infected tissue (leaf pieces of ~6 cm length), 8–9 days after infection (or 1 day before sporulation) (see Note 1).

2. To remove external contaminating organisms, wash the tissue with tap water several times, then incubate for 3 min in 2 % sodium hypochlorite. Wash with tap water three times or until sodium hypochlorite is completely removed, then incubate the tissue for 1 min in chilled 70 % ethanol. Wash tissue with chilled MilliQ water several times to remove the alcohol. Dry the tissue as much as possible with sterile paper towel.

3. Using a Waring blender, homogenize the infected plant material in 150 ml of homogenization buffer at maximum speed for 25 s.

4. Filter the homogenate through a 100 μm nylon mesh by gravity flow. Recover the solid particles retained on the mesh and return them to the blender, add the remaining 100 ml of homogenization buffer and blend for 15 s at 18,000 rpm. Pass through the 100 μm mesh to remove cell debris and combine the two filtrates.

5. Pass the filtrate (~250 ml) through the 20 μm mesh by gravity flow and distribute the new filtrate in six chilled Oak Ridge round-bottom centrifuge tubes, 45 ml capacity (see Note 2).

6. Centrifuge the tubes at 1,080 × g for 15 min at 4 °C, using a centrifuge with swingout rotor. In the meantime, prepare Percoll-30 solution: 24 ml Percoll, 30 ml MilliQ water, and 6 ml 10× IB.

7. Resuspend each pellet very gently in 2 ml of ice-cold 1× IB using a 1 ml pipette; combine all of the resuspended pellets and bring the total volume to 20 ml with 1× IB. Mix the resuspended pellets (20 ml) with the Percoll-30 solution to a total volume of 80 ml. Split the mixture into four 45 ml polycarbonate round-bottom tubes.

8. Centrifuge at 25,000 × g for 30 min at 4 °C without braking.

9. Carefully remove the tubes from the centrifuge, and draw off the first 10 ml from each tube very slowly using a 10 ml pipette and an automatic pipette controller under low suction speed (see Note 4). Dilute the pooled 40 ml haustorial fraction 1:10
into 360 ml of 1× IB. Mix well and distribute the mixture into ten chilled 45 ml Oak Ridge round-bottom centrifuge tubes.

10. Centrifuge the tubes at 1080 × g for 15 min at 4 °C, using a centrifuge with swingout rotor (steps 11–13 are optional, see Note 5). If following steps 11–13 resuspend the pellets in 10 ml of 1× IB, otherwise go to step 14.

11. Prepare Percoll-25 solution: 10 ml Percoll, 17 ml MilliQ water, and 3 ml 10× IB. Mix this solution with the resuspended pellets for a final volume of 40 ml.

12. Centrifuge at 25,000 × g for 30 min at 4 °C without braking in two 45 ml polycarbonate round-bottom tubes. Draw off the first 10 ml from each tube very slowly using a 10 ml pipette and an automatic pipette controller under low suction speed. Dilute the pooled haustorial fractions 1:10 in 180 ml of 1× IB. Mix well and distribute this mixture into chilled 45 ml Oak Ridge round-bottom centrifuge tubes.

13. Centrifuge the tubes at 1080 × g for 15 min at 4 °C, using a centrifuge with swingout rotor.

14. Resuspend the pellets in 1× IB to a final volume of 4 ml, and transfer the suspension to a glass vial with plastic lid, 6 ml capacity.

15. Add 200 μl of 1 mg/ml Con A-Alexa 488 (Invitrogen), cover the container with foil, and mix gently on a rotary mixer for 20 min at room temperature (or for 45 min at 4 °C).

16. Pellet the haustoria at 4,000 × g in a benchtop centrifuge for 5 min at 4 °C. Remove the supernatant and wash the pellet twice with 1× IB. Resuspend in a final volume of 4 ml of 1× IB; keep sample on ice in the dark and proceed directly to FACS sorting.

### 3.2 Flow Cytometry of Pre-isolated Haustoria

This section describes the use of flow cytometry to sort haustoria based on their fluorescence after labeling with Con A-Alexa 488. The Alexa 488 fluor is excited using a 13 mW 488 nm solid-state laser (Sapphire, Coherent Inc. Santa Clara, CA) and detected using a 502 nm longpass as well as a 530/30 nm bandpass filter. Chlorophyll autofluorescence is excited with the same 488 nm laser and detected using a 655 longpass and 695/40 bandpass filter.

1. Dilute the haustorial sample 1:10 with chilled 1× PBS to adjust the concentration of the chloroplast–fungal cell mixture to approximately 2.5–5.0 × 10⁶ particles/ml.

2. Set up and optimize the cell sorter (see Note 6). For fungal haustoria (5–20 μm in diameter): 100 μM nozzle, 20 psi sheath pressure.
3. Use 1× PBS as sheath fluid, run cells using the lowest sample pressure with a resulting sample rate of 5,000–10,000 events per second. Collect sorted cells in 5 ml glass tubes containing 1× PBS, keeping them at 4 °C, and agitating them periodically at 300 rpm to prevent settling.

4. Set fluorescence and scatter parameters. Identify and exclude chlorophyll-containing particles based on chlorophyll autofluorescence. Identify and select the population of Con A-positive particles by Alexa 488 fluorescence (see Fig. 1a).

Fig. 1 Purification of *Puccinia striiformis* f.sp. *tritici* haustoria by FACS, and microscopic analysis of pre- and post-sorted samples. Haustoria were partially purified by Percoll gradients as described, then stained with Con A-Alexa 488 prior to sorting. (a) Flow-cytometric analysis of pre-isolated haustoria showing the discrimination of two populations based on red chloroplast autofluorescence and green Alexa 488 fluorescence from labeled haustoria. (b) Flow-cytometric analysis of a sorted sample to test the efficiency of the sorting process, average purity of 98 %. (c) Microscopic image of a sample prior to sorting, with a haustoria:chloroplast ratio of about 1:80. (d) Microscopic image of a sample post-sorting, with a haustoria:chloroplast ratio of about 49:1. In (a) and (b) plots displaying flow cytometry data, haustoria are represented by the green dots and chloroplasts by the blue dots.
Use forward and side scatter information to identify and exclude doublets by comparing the forward scatter height and width and the side scatter height and width. To optimize the forward scatter signal, use a neutral density 2 filter in front of the forward scatter detector.

5. Confirm sort purity by running a small amount of the sorted cells to determine the percentage of positive events for the sort parameters (Fig. 1b) (see Notes 7 and 8).

6. Pool the samples from FACS purification and centrifuge them at $1,080 \times g$ for 15 min at 4 °C. Discard the supernatant and snap-freeze the final pellet in liquid nitrogen, and store it at −80 °C until required.

4 Notes

1. Harvest tissue as late as possible before sporulation to increase the number of haustoria; however, it should be done at least 1 day before pustule eruption. In wheat plants infected with *P. striiformis* and maintained at 18 °C post-infection, this time is usually 8–9 days.

2. The pore size of the meshes is critical for reducing contaminating particles before FACS. Different fungal species display a diverse range of haustoria size, thus small-scale isolations to test different pore sizes and microscopic analyses are necessary to optimize the filtration step.

3. The stock solution was prepared as per manufacturer’s instructions but without addition of sodium azide.

4. Haustoria from different fungal species exhibit slightly different density characteristics. It is highly recommended to establish the density level at which the majority of haustoria resolve. Small-scale density tests can be performed by following the protocol above until step 8, Subheading 3.1. Once the tubes are removed from the centrifuge, aliquots of 1 ml are taken carefully with a wide-bore pipette tip, and the number of haustoria can be counted in each aliquot using a hemocytometer.

5. Steps 11 and 12 are included specifically to reduce the complexity of the sample prior to flow cytometry. However, depending on the final use of the sorted haustoria, these two steps can be omitted.

6. The procedure of setting up a flow cytometer varies depending on the machine and needs to be performed by appropriately trained personnel.

7. Typically, $9 \times 10^5$–1.5 × 10^6 purified haustorial cells can be purified from 15 to 20 g of wheat tissue heavily infected with *P. striiformis*. 
8. Vital tests should be applied at this stage to verify the viability of haustoria post-sorting. A high percentage of viable cells could be very important for some subsequent applications as isolation of intact RNA or proteins. In our studies we used CellTracker™ Orange CMRA (Invitrogen, Catalog number C34551) successfully following the manufacturer’s instructions, and obtained ≥98% viable haustoria post-sorting.

References