# Genomic resources to study virulence and evolution of cereal rust fungi

A THESIS SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA BY

Eva Celeste Henningsen

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE

Professor Brian Steffenson

May 2021

# Acknowledgements

Thank you to the staff at the Minnesota Supercomputing Institute at the University of Minnesota for technical assistance, particularly Nick Dunn who helped me again and again with software installation and troubleshooting. Thank you also to Jana Sperschneider who was extremely supportive during my first time running a genome assembly pipeline. Without their help, none of this work would have been possible.

I would also like to thank my scientific and personal mentors: Melania Figueroa, for her unwavering support through difficult times and role model as a successful female scientist and leader; Brian Steffenson, for his continued support throughout my career and for entrusting to me important research; and Feng Li, my science big sister, who guided me through the beginning of my journey and remains a source of inspiration to me today.

# Dedication

I dedicate this thesis to my parents and my cat Jasmine.

# Abstract

Stem rust caused by *Puccinia graminis* f. sp. *tritici* (*Pgt*) and crown rust caused by *Puccinia coronata* f. sp. *avenae* (*Pca*) are global threats the production of wheat and oat, respectively. Fast evolving populations of both *Pgt* and *Pca* limit the efficacy of plant genetic resistance and constrain disease management strategies. Chapter 1 provides background information about both rust fungi and their biology, shares a comprehensive review of the available genome resources in the rusts, and highlights some advancements in rust research and how they can be utilized. Chapter 2 describes a study where my colleagues and I developed a pipeline for identifying candidate susceptibility genes for future study of stem rust virulence using comparative transcriptome-based and orthology-guided approaches. The analysis was targeted to genes with differential expression in *T. aestivum* and genes suppressed or not affected in *B. distachyon* and reports several processes potentially linked to susceptibility to *Pgt*, such as cell death suppression and impairment of photosynthesis. The approach was complemented with a gene co-expression network analysis to identify wheat targets to deliver resistance to *Pgt* through removal or modification of putative susceptibility genes. This work could help further the understanding of the molecular mechanisms that lead to rust infection and disease susceptibility; this in turn could deliver novel strategies to deploy crop resistance through genetic loss of disease susceptibility. A significant contribution of this work is a pipeline that can be adapted to study virulence of other rust fungi. Finally, Chapter 3 describes a high-quality genome assembly of *Pca* isolate 203. The ultimate goal of the assembly is to provide the first fully haplotype-phased, chromosome level reference for *Pca*. To this end, PacBio long reads and Illumina short reads were obtained to create the initial draft assembly, while Hi-C reads were collected to order contigs and phase the genome. Contigs were assigned to haplotype bins using gene synteny initially, and these bins were aligned to the *Pgt* 21-0 A haplotype genome to evaluate the probable number of chromosomes and possible chromosome sizes. Future steps for completing the high-quality assembly include an iterative process to fix haplotype phase swaps through manual curation and scaffolding, final chromosome assignment, and annotation with RNAseq data. A collection of publications with my contributions is provided in the appendix section.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1: Introduction

## 1.1 Overview of wheat stem rust and oat crown rust diseases

The stem rust fungus and crown rust fungus are two of the most destructive pathogens of wheat and oat, respectively, as epidemics can severely affect production (Singh et al. 2011; Nazareno et al. 2018). Rusts are basidiomycete fungi of the order Puccinales, with many belonging to the genus *Puccinia* (Aime et al. 2017; Kolmer et al. 2018). Rust fungi have a high degree of variation in their life cycles, with some being macrocyclic, demicyclic, or microcyclic (Petersen 1974; Lorrain et al. 2019). In addition, some rusts are autoecious and complete their full life cycle on one host, while others, including important agricultural pathogens, are heteroecious and require multiple hosts to complete the full life cycle (Petersen 1974; Lorrain et al. 2019). The asexual cycle is completed on the telial host; urediniospores are produced on and re-infect telial host plants when weather conditions are favorable. The sexual cycle of heteroecious rusts is completed when basidiospores, which germinate from teliospores, infect the aecial host, and develop into pycnia. These produce pycniospores which, in fertilizing compatible flexuous hyphae, cause the development of aecia on the abaxial leaf face. The aecia produce aeciospores and finally, the aeciospores infect the grass host and the cycle is completed. Microcyclic rusts are by necessity autoecious, and can only complete the asexual cycle, though there are alternative methods of generating genetic diversity which some microcyclic rusts have been observed to possess (Ono 2002). Demicyclic rusts only complete a sexual cycle, as they lack urediniospores. Both the asexual and sexual cycles are present in macrocyclic rusts, which

possess all five spore stages and can be either autoecious or heteroecious (Petersen 1974; Lorrain et al. 2019).

Wheat stem rust, caused by *Puccinia graminis* f. sp. *tritici* (*Pgt*) is one of the most devastating wheat diseases worldwide at present, although its importance in the United States before and into the mid-20$^{th}$ century was greater than today (Roelfs 1985; Singh et al. 2011). *Pgt* is heteroecious, relying on wheat (*Triticum spp.*) as the telial host and barberry (*Berberis spp*.) as the aecial host (Roelfs 1985). The presence of barberry generally increases infection on wheat by providing an early and local source of spring inoculum and also acting as the locus of sexual recombination (Roelfs 1985; Olivera et al. 2019). Common barberry (*Berberis vulgaris*) was introduced to North America by Europeans in the 17$^{th}$ century and was used for medicine, fruit production, and as wind breaks adjacent to wheat fields (Peterson 2003). The close proximity of the telial and aecial hosts created a perfect environment for large, diverse populations of the stem rust pathogen, which resulted in frequent epidemics (Roelfs 1985). After a particularly severe epidemic in 1916, a barberry eradication program began in the United States, with federal and state governments hiring labor to locate and remove barberry plants from 1918 to 1980 (Peterson 2001b; Figueroa et al. 2016). The eradication program not only removed the early spring inoculum source of *Pgt*, but also eliminated the potential for sexual recombination to occur between individuals. This resulted in a dramatic drop in the number of stem rust races reported in the United States each year during the annual survey; in 2014, only one race (QFCSC) was found at all sampling locations (2014 Wheat Stem Rust Observations in the U.S. 2014).

Despite the extreme homogeneity of *Pgt* in North America, the emergence of the highly virulent Ug99 isolate in Africa highlights the diversity evident elsewhere in the world and represents a serious threat to food security (Pretorius et al. 2000). Ug99 (race TTKSK) and other derivative isolates were found to be virulent on 85-95% of tested wheat varieties and more than 95% of tested barley varieties, prompting efforts to identify and integrate effective resistance genes into wheat and barley germplasm (Singh et al. 2011; Steffenson et al. 2017). Recent research into the origin of Ug99 revealed that somatic hybridization and whole nuclear exchange were most likely responsible for the emergence of this highly divergent lineage (Li et al. 2019*).

A similarly destructive rust pathogen of oats is *Puccinia coronata* f. sp. *avenae* (*Pca*), which causes oat crown rust. *Pca* is also heteroecious, but oat (*Avena spp.*) is the telial host and buckthorn (*Rhamnus spp.*) is the aecial host (Simons 1985; Nazareno et al. 2018). Like with *Pgt*, infection is generally increased in areas with high buckthorn prevalence due to the spring inoculum source and genetic diversity generated from sexual recombination (Simons 1985; Miller et al. 2021*). Buckthorn, like barberry, was introduced by Europeans some time during the 19$^{th}$ century for its desirable medicinal and later ornamental properties; however, buckthorn turned out to be highly aggressive and has invaded all suitable environments in the United States and into Canada (Knight et al. 2007; Kurylo and Endress 2012). A highly diverse *Pca* population in North America is maintained on buckthorn and wild oats, from which genotypes that overcome disease resistance are selected for when single-gene oat varieties are released. This results in a rapid boom and bust cycle where successful pathogen races amplify to epidemic levels following the release of a new single-gene cultivar (Figure 1.1) (Simons 1985; Mundt 2014). Despite the

negative ecosystem impacts of buckthorn and the connection between buckthorn and *Pca* diversity and virulence, large-scale programs to eradicate the invasive have not occurred and current efforts are small-scale and voluntary. A recent study using a population genomics approach revealed that both sexual and asexual propagation of *Pca* contribute to population diversity and therefore to evolution of virulence (Miller et al. 2021*). Thus, eradication of buckthorn may only reduce the number of genotypes that become prevalent in specific geographic regions. Other processes such as mutation or somatic hybridization in clonal populations can also alter the population dynamics of this pathogen, as they do in other rust species (Figueroa et al. 2020*).

### 1.2 Biology of *Puccinia graminis* f. sp. *tritici* and *Puccinia coronata* f. sp. *avenae*

Both *Pgt* and *Pca* are macrocyclic, heteroecious fungi that have similar life cycles and infection processes (Petersen 1974). Urediniospores land on a leaf of a susceptible grass host, imbibe water, and germinate. The germ tube, likely via thigmotropism, locates a stomata on the host leaf and forms an appressorium above it. Spore germination is inhibited by light, and germ tubes avoid light while growing (Staples and Macko 1984). The appressorium produces a penetration peg, which enters the stomata and allows the fungus to continue growing within the host leaf. Infection hyphae form haustorial mother cells next to host plant mesophyll cells, and haustorial formation is initiated. The haustoria is the interface between the fungus and the host, and is where nutrients and effectors are exchanged (Bozkurt and Kamoun 2020). A compatible interaction is completed when erumpent pustules form where urediniospores are released.

Each step in the infection process represents a point of interaction between the host and pathogen that determine whether the ultimate outcome is a successful infection or not. Recognition of rust fungi in plants is mediated by an innate immunity system that activates defenses to confer disease resistance. In this system, dominant resistance (R) genes in the host plant confer recognition of specific Avirulence (Avr) genes from the pathogen. R genes often encode intracellular immune receptor proteins belonging to the nucleotide binding leucine rich repeat receptor (NLR) class (Dodds and Rathjen 2010). These receptors recognize pathogen 'effector' proteins that are delivered into host cells during infection to suppress host basal defenses and facilitate infection (Periyannan et al. 2017). This type of resistance is known as race-specific resistance, as it is effective to combinations of Avr factors (effectors) that define a specific race of the pathogen. Effective R genes in a crop exert strong selection on the pathogen to evolve and escape recognition, therefore variation in effector gene content and sequence can be responsible for the emergence of new virulence traits that can overcome deployed R genes in cultivars (Figueroa et al. 2020*; Henningsen et al. 2020*). This gene-for-gene interaction has been the pillar for breeding disease resistant crop varieties and represents the underlying model of host-pathogen co-evolution (Ellis et al. 2014). In the oat crown rust system, about 100 R genes (*Pc* genes) have been proposed based on genetic data (Admassu-Yimer et al. 2018). Unfortunately, most of them have been overcome by pathogen evolution as they were often released singly in oat varieties. Adding to this, none of the Avr factors in *Pca* have been identified; in fact, only three Avr effector genes have been isolated for any cereal rust fungal species (Saintenac et al. 2013; Chen et al. 2017; Upadhyaya et al., in press). The

cloning of effector genes from rust fungi has been difficult due the genomic complexity of these organisms (Figueroa et al. 2020*).

The full life cycles of *Pgt* and *Pca* can be summarized as follows: urediniospores are produced on the telial host in uredinia until environmental conditions change to favor production of telia from the same uredium. The telia will overwinter/oversummer until favorable conditions occur, at which point they germinate and produce basidiospores. Basidiospores infect the aecial host and from the infection produce pycnia of two different mating types. These pycnia produce pycniospores, which fertilize the flexuous hyphae of the opposite mating type. The fertilized pycnia will produce hyphae that grow through the leaf tissue and produce aecia on the abaxial leaf surface. Aeciospores are produced in the aecia and infect the telial host.

### 1.3 Genomics of rust fungi

Like many rust fungi, *Pgt* and *Pca* are dikaryotic for the duration of their life cycle on the telial host, so they can be treated as diploids in genetic studies (Roelfs 1985; Simons 1985). However, due to their relatively large genome size (~90 Mb/haplotype for *Pgt* ~100 Mb/haplotype for *Pca*), high repeat content, and recalcitrance to culturing, genomic studies have lagged behind other plant pathogenic fungi (Williams 1984; Tavares et al. 2014; Aime et al. 2017; Figueroa et al. 2020). The rust species *Pgt* and *Puccinia striiformis* f. sp. *tritici* (wheat stripe rust fungus) arguably have the best genomic resources at present, as both have two fully-phased genome assemblies (Schwessinger et al. 2018; Xia et al. 2018); other rust genome assemblies are represented with hundreds to thousands of contigs, and most of these are not haplotype phased (Table 1.1) (Duplessis et al. 2011; Cuomo et al. 2017;

Porto et al. 2019; Wu et al. 2020). There are 18 chromosomes in *Pgt* that have a dikaryotic size of 170 Mbp with about 34% of the genome consisting of gene space (Li et al. 2019*). *Pca* does not have a chromosome-level assembly, but flow cytometry experiments placed haploid chromosome number between 16-20 (Boehm 1992). The current partially-phased assemblies are 150.47 and 166.28 Mbp for the US-derived isolates 12SD80 and 12NC29, respectively (Miller et al. 2018); however, the primary contigs of each total to between 99 Mbp and 105 Mbp, suggesting a diploid genome size of about 200 to 210 Mbp (Miller et al. 2018). Currently there are 5 genome assemblies for *Pgt* available in the National Center for Biotechnology Information (NCBI) database, representing isolates CRL 75-36-700-3 (88.72 Mbp), RKQQC (145.10 Mbp), Pgt21-0 (176.85 Mbp), Ug99 (176.24 Mbp), and UK-01 (164.29 Mbp) released between 2007 and 2020 (Duplessis et al. 2011; Rutter et al. 2017; Li et al. 2019*; Lewis et al. 2018). A website for wheat rust genomic resources was recently launched and can be found at https://www.wheat-rust-genomics.com/. The only genome assemblies for *Pca* on NCBI are 12SD80 and 12NC29 as described above (Miller et al. 2018). A summary of existing rust genomes as adapted from a previous publication (Figueroa et al. 2020*) is provided in Table 1.1.

### 1.4 Advances in rust research driven by genomic resources

While genetics work in rust fungi was pioneering and revolutionary at one time with the development of plant immunity concepts like the gene-for-gene hypothesis (Flor 1971), the complexity of the rusts and their biotrophic nature means that only now do we have the tools to generate accurate, phased genome references and study genetic diversity without the need for biparental populations. The lag in genomic resources has also led to a lag in the discovery of effectors, which in other systems such as *Pseudomonas syringae*

(bacteria), *Phytopthora* spp. and *Pythium* spp. (oomycetes), and *Magnaporthe oryzae* (fungus), effector knowledge is already robust (Wang and Jiao 2019; Ai et al. 2020; Laflamme et al. 2020; Petit-Houdenot et al. 2020). With the recent generation of high-quality reference genomes and the establishment of pipelines to produce more references, new research opportunities have emerged in identifying and cloning effectors in rust fungi. So far, three effectors have been cloned in rust fungi and these are *AvrSr50*, *AvrSr35*, and *AvrSr27* (Chen et al. 2017; Salcedo et al. 2017; Figueroa et al, in press*; Upadhyaya et al. in press). Pioneering effector research in rust fungi was conducted in the flax-flax rust system (*Melampsora lini – Linum usitatissimum*), in which 19 flax resistance genes have been cloned and several corresponding *Avr* genes in the flax rust pathogen have been identified; so far the interactions between the R and *Avr* genes in rusts have been shown to be direct (Figueroa et al. 2020*), similar to findings in other biotrophic fungi such as powdery mildew (Saur et al. 2019).

### 1.5 Aims of this thesis

The work presented in my thesis was driven by the current needs to develop genomic resources for hypothesis-driven studies to determine virulence factors and mechanisms important for rust fungi. As much as there is a need to study the pathogen, there are still many questions about rust susceptibility determinants in the host. Chapter 2 aims to harness the resources developed in systems with less genetic complexity than that of hexaploid wheat, namely *Arabidopsis thaliana* and *Hordeum vulgare*, to advance our understanding of rust virulence. Susceptibility factors have been identified in both *A. thaliana* and *H. vulgare* which may be relevant to wheat as well; however, the difficulty of functional validation in wheat necessitates a focused approach to testing candidates.

We took a comparative RNA-seq approach to determine strong candidates for susceptibility genes in wheat by including near-isogenic wheat lines W2691 (susceptible) and W2691+*Sr9b* (resistant), as well as *Brachypodium distachyon* accession Bd21-3. We then identified the orthologs of known susceptibility genes in *A. thaliana* and *H. vulgare* in wheat and *B. distachyon* and determined what genes were co-regulated with strong candidates. Ultimately, the goal of the project was to develop a pipeline to harness the previously developed resources from highly studied model organisms for use in more complicated plant-pathogen systems.

The primary goal of Chapter 3 is to generate a high-quality genome reference for the historic *Pca* isolate 203. Previous genome references for *Pca* only have partial haplotype phasing and are not chromosome-level, but new opportunities for generating a better *Pca* assembly arrived with the development of better resources for genome assembly in non-haploid organisms. Isolate 203 was considered for sequencing as a reference due to the likelihood that it carries more *Avr* effectors than contemporary isolates of *Pca*. Isolate 203 is also connected to the devastating epidemics of *Cochliobolus victoriae* on Victoria oats in the U.S. back in the 1940's; Victoria was released to provide resistance to 203, but the *C. victoriae* epidemics occurred as a result of widespread planting of Victoria. The high-quality reference for 203 is expected to further the understanding of rust genomics and genetics.

### 1.6 Contributions to publications cited in the introduction

Although beyond the scope of this thesis, I have contributed to various peer-reviewed publications that represent significant advances to the field of molecular plant pathology.

Below I list these publications and my role in the completion of the work. The abstracts of these publications have been added in the appendix section and are also noted with asterisks (*) when cited throughout Chapter 1.

Appendix A: Li et al. (2019). Emergence of the Ug99 lineage of the wheat stem rust pathogen through somatic hybridization. Nature Communications. 10:5068

- Responsible for orthology analyses during the haplotype assignment steps and manuscript revisions

Appendix B: Miller et al. (2021). Increased virulence of *Puccinia coronata* f. sp. *avenae* populations through allele frequency changes at multiple putative *Avr* loci. PLoS Genetics. 16(12)

- Involved with the interpretation of GWAS results and manuscript revisions

Appendix C: Figueroa et al. (2020). Evolution of virulence in rust fungi — multiple solutions to one problem. Current Opinion in Plant Biology. 56

- Contributed to writing, editing, and revising the manuscript draft

Appendix D: Henningsen et al. (2020). *Rpg7*: A New Gene for Stem Rust Resistance from *Hordeum vulgare* ssp. *spontaneum*. Phytopathology. 111(3)

- Conducted the inoculation experiments, scoring, and managed data entry for $F_2$, $F_3$, and allelism experiments
- Filtered GBS marker data, created linkage and QTL maps, and generated plots
- Identified nearby candidate genes from QTL results

- Wrote, edited, submitted, and revised manuscript

Appendix E: Figueroa et al. (2021) Tactics of host manipulation by intracellular effectors from plant pathogenic fungi. Current Opinion in Plant Biology. 62

- Contributed to writing, editing, and revision of the manuscript draft

**Table 1.1**. Available rust genome references (adapted from Figueroa et al. 2020).

| Species | Host | Haploid genome size | Haplotype-phased? | Reference(s) |
|---|---|---|---|---|
| *Puccinia striiformis* f. sp. *tritici* | Wheat | 74-83 Mbp | Yes (2 isolates) | (Cantu et al. 2011, 2013; Zheng et al. 2013; Schwessinger et al. 2018; Xia et al. 2018; Schwessinger et al. 2020) |
| *Puccinia graminis* f. sp. *tritici* | Wheat, barley | 88 Mbp | Yes (2 isolates) | (Duplessis et al. 2011; Upadhyaya et al. 2015; Li et al. 2019) |
| *Puccinia striiformis* f. sp. *hordei* | Barley | 89 Mbp | No | (Xia et al. 2018) |

| | | | | |
|---|---|---|---|---|
| *Puccinia coronata* f. sp. *avenae* | Oat | 100 Mbp | Yes (2 isolates partially) | (Miller et al. 2018) |
| *Melampsora larici-populina* | Poplar | 101 Mbp | No | (Duplessis et al. 2011) |
| *Puccinia triticina* | Wheat | 135 Mbp | No | (Cuomo et al. 2017; Wu et al. 2020) |
| *Puccinia hordei* | Barley | 150 Mbp | No | (Chen et al. 2019) |
| *Melampsora lini* | Flax | 240 Mbp | No | (Nemri et al. 2014) |
| *Hemileia vastatrix* | Coffee | 550 Mbp | No | (Porto et al. 2019) |
| *Phakopsora pachyrhizi* | Soybean | 1.06 Gbp | No | https://mycocosm.jgi.doe.gov/Phapa1 |

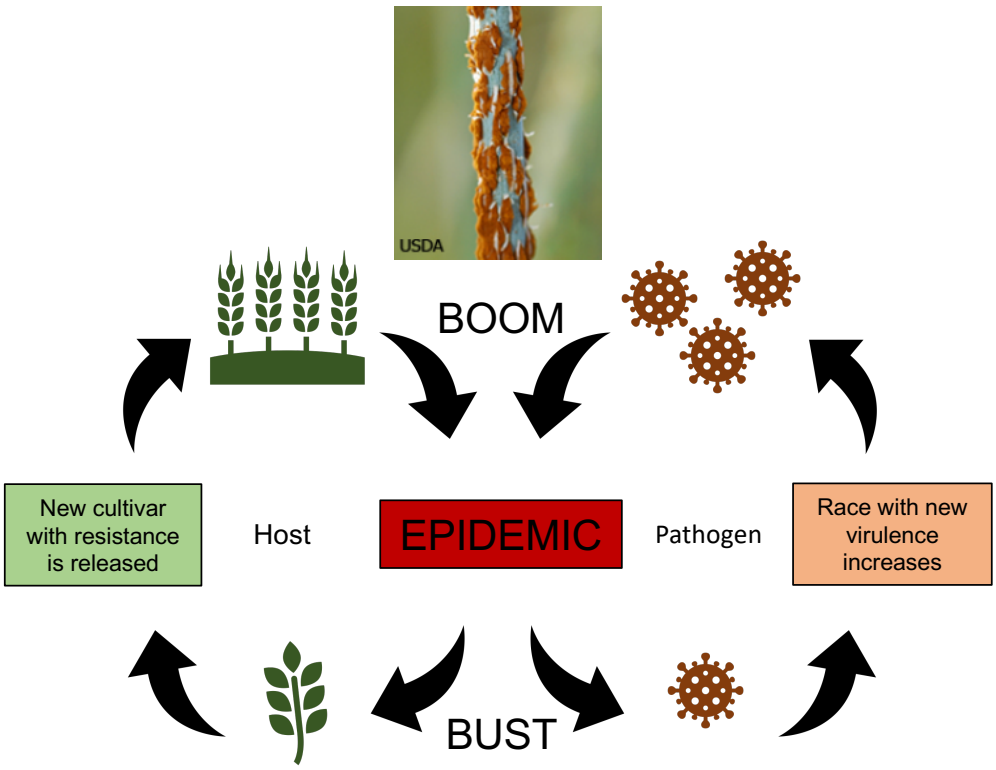



**Figure 1.1**. Diagram of the boom-and-bust cycle frequently observed in *Pgt* and *Pca*. As a new cultivar with single-gene resistance is released and widely planted, only individuals with virulence to the new resistance gene survive and reproduce due to high selection pressure. Once the virulent pathogen race increases enough, an epidemic occurs and the area sown to the once-effective resistant cultivar diminishes. The cycle repeats when a new single-gene cultivar is introduced. Photo of stem rust infection from the Cereal Disease Laboratory, USDA-ARS.

# Chapter 2: Identification of candidate susceptibility genes to *Puccinia graminis* f. sp. *tritici* in wheat[1]

**Contributions to Chapter 2**

This chapter is published in Frontiers in Plant Science. I was responsible for:

- Mapping RNAseq data to the wheat and *Brachypodium* genomes and calling differentially expressed genes for inoculated and mock treatments
- Orthogroup analysis for wheat, *Brachypodium*, *Arabidopsis*, and barley genes
- GO term enrichment analysis for all expressed genes
- GO term enrichment analysis for networks
- Creation of GO enrichment, network, and differential expression figures
- Creation of main text and supplemental tables
- Writing, editing, and revision of manuscript

---

[1]This work is published in *Frontiers in Plant Science*.

As stated in the manuscript:

Author contributions: Melania Figueroa, Cory D. Hirsch, Chad Meyers, Shahryar F. Kianian conceived and designed the study; Vahid Omidvar, Marisa E. Miller and Feng Li conducted the experiments; Eva C. Henningsen, Rafael Della Coletta, Vahid Omidvar, Erin Gilbert, Jean Michno, Cory D. Hirsch, Sean P. Gordon, John P. Vogel and Marisa E. Miller contributed to data analysis. Eva C. Henningsen, Brian J. Steffenson, Melania Figueroa, Shahryar F. Kianian interpreted results. Eva C. Henningsen, Vahid Omidvar, Melania Figueroa, and Cory D. Hirsch wrote the manuscript; all authors contributed to manuscript editing and revisions and approved the submitted version.

## 2.1 Introduction

Stem rust caused by *Puccinia graminis* f. sp. *tritici* (*Pgt*) is one of the most devastating foliar diseases of wheat (*Triticum aestivum*) and barley (*Hordeum vulgare*). The economic relevance of this pathogen to food security is demonstrated by the impact of historical and recent epidemics (Pretorius et al. 2000; Peterson 2001a; Olivera et al. 2015; Singh et al. 2015; Bhattacharya 2017; Steffenson et al. 2017). Consistent with its biotrophic lifestyle, *Pgt* develops an intricate relationship with its host in order to acquire nutrients and survive. Early stages of infection involve the germination of urediniospores (asexual spores) and host penetration through the formation of appressoria over stomata (Staples and Macko 1984). As the fungus reaches the mesophyll cavity of the plant, it develops infection hyphae which penetrate plant cell walls and differentiate into specialized feeding structures, known as haustoria. Haustorial development takes place during the first 24 hours post-infection and is critical for colony establishment and sporulation that re-initiates the infection cycle (Harder and Chong 1984). Similar to other plant pathogens, cereal rust infections involve the translocation of effectors to the plant cell as a mechanism to shut down basal defenses activated by PAMP-triggered immunity (PTI) and manipulate host metabolism (Dodds and Rathjen 2010; Couto and Zipfel 2016). In rust fungi, the haustorium mediates the secretion of effectors, although the underlying molecular mechanism that facilitates this process is not known (Garnica et al. 2014; Petre et al. 2014). The plant targets of effectors and other plant genes that mediate compatibility and facilitate pathogen infection are often regarded as susceptibility (*S*) genes (Lapin and Van den Ackerveken 2013; van Schie and Takken 2014; Lo Presti et al. 2015).

To avoid infection by adapted pathogens, plants employ effector-triggered immunity (ETI) which is mediated by the recognition of effectors by nucleotide-binding domain leucine-rich repeat (NLR) receptors (Flor 1971; Dodds and Rathjen 2010; Garnica et al. 2014; Petre et al. 2014). These specific recognition events often induce localized cell death at infection sites (hypersensitive response, HR) which restrict pathogen growth. In wheat-rust interactions, ETI is manifested by the reduction or absence of fungal growth and sporulation (Periyannan et al. 2017). The use of NLR genes to provide crop protection was a critical component of the Green Revolution which diminished the impact of stem rust epidemics (Ellis et al. 2014). While this approach still contributes to the development of wheat cultivars with genetic resistance to stem rust, the durability of such resistant cultivars is hampered by the evolution of rust populations to avoid recognition by NLRs. Given the economic and environmental advantages of genetic disease control over chemical applications, the identification of alternative genetic sources of resistance are a priority to secure future wheat production. In this context, the discovery of *S* genes could have important translational applications for agriculture and potential durable disease control. Mutations in *S* genes, although often recessive, could shift a genotype to a non-suitable host due to alterations in initial recognition stages or loss of pathogen establishment requirements (van Schie and Takken 2014; Lo Presti et al. 2015).

The genetic factors that contribute to wheat susceptibility to biotrophic pathogens such as rust fungi remain largely unknown. Numerous structural and physiological alterations have been observed in wheat-rust compatible interactions. At early infection stages, 4-6 days post-inoculation (dpi), the cytoplasm of infected mesophyll cells increases in volume and an extensive network of the endoplasmic reticulum is built near the haustorium (Bushnell

1984). The nucleus of infected cells also increases in size and migrates towards the haustorium, and in some cases both structures appear in proximity. These observations suggest that plant cells undergo a massive transcriptional reprogramming to either accommodate rust colonization or initiate a cascade of plant defenses to prevent infection. In addition, many biotrophs are known to increase the ploidy of host cell nuclei near infection sites (Wildermuth 2010). Advances in next generation sequencing and data mining bring new opportunities to deepen our understanding of plant-pathogen interactions and the relationship between plant metabolism and disease resistance or susceptibility. Several transcriptome profiling studies comparing compatible and incompatible wheat-rust interactions provide strong evidence for the complexity of these interactions (Bozkurt et al. 2010; Zhang et al. 2014; Chandra et al. 2016; Dobon et al. 2016; Yadav et al. 2016). Although *S* genes in rust pathosystems are largely unknown, several susceptibility factors to other plant pathogenic fungi have been identified in *Arabidopsis thaliana*, *Hordeum vulgare* (barley), and solanaceous plants (van Schie and Takken 2014; Zaidi et al. 2018).

To expand our knowledge of wheat-rust interactions and identify candidate *S* genes to direct future functional studies, we conducted a comparative RNA-seq analysis of the molecular responses to *Pgt* in compatible and incompatible interactions. We included a susceptible genotype ('W2691') of *Triticum aestivum* (bread wheat) and the near-isogenic line (NIL) 'W2691+*Sr9b*' containing the resistance gene *Sr9b*, which confers race-specific responses to various *Pgt* isolates (McIntosh et al. 1995). We also included the related grass species *Brachypodium distachyon*, which is recognized as a non-host to various cereal rust species (Kellogg 2001; Ayliffe et al. 2013; Figueroa et al. 2013, 2015; Bettgenhaeuser et al. 2014, 2018; Gilbert et al. 2018; Omidvar et al. 2018). As part of our analysis, we

examined the expression profiles of *T. aestivum* and *B. distachyon* orthologs of several known *S* genes in *Arabidopsis thaliana*, *H. vulgare*, as well as other characterized *S* genes in *T. aestivum*, and identified groups of genes co-regulated with these *S* gene candidates. In conclusion, this study provides an overview of global expression changes associated with failure or progression of *Pgt* infection in *T. aestivum* and *B. distachyon* and insights into the molecular processes that define disease incompatibility.

## 2.2 Materials and Methods

### 2.2.1 Plant and fungal materials

Two near-isogenic lines of *T. aestivum*, 'W2691' (Luig and Watson 1972) and 'W2691' carrying the *Sr9b* gene (referred to onward as 'W2691+*Sr9b*', U.S. National Plant Germplasm System Accession Identifier: CItr 17386) and the *B. distachyon* Bd21-3 inbred line (Vogel and Hill 2008) were used in this study. *T. aestivum* and *B. distachyon* seeds were received from the USDA-ARS Cereal Disease Laboratory (CDL) St. Paul, MN, USA and the USDA-ARS Plant Science Unit, St. Paul, MN, USA, respectively. The fungal isolate *Puccinia graminis* f. sp. *tritici (Pgt)* (isolate # CDL 75-36-700-3 race SCCL) (Duplessis et al. 2011) was obtained from the USDA-ARS CDL.

### 2.2.2 *Pgt* infection of *T. aestivum* and *B. distachyon* genotypes

*B. distachyon* seeds were placed in petri dishes with wet grade 413 filter paper (VWR International) at 4°C for five days and germinated at room temperature for three days before sowing to synchronize growth with wheat plants which did not require stratification. Seeds of both wheat and *B. distachyon* were sown in Fafard® Germination Mix soil (Sun

Gro Horticulture, Agawam, MA, U.S.A.). All plants were grown in growth chambers with a 18/6 hour light/dark cycle at 21/18°C light/dark and 50% relative humidity. Urediniospores of *Pgt* were activated by heat-shock treatment at 45°C for 15 minutes and suspended in Isopar M oil (ExxonMobil) at 10 mg/mL concentration. Inoculation treatments consisted of 50 µl of spore suspension per plant, whereas mock treatments consisted of 50 µl of oil per plant. Fungal and mock inoculations were conducted on seven-day old wheat plants (first-leaf stage) and twelve-day old *B. distachyon* plants (three-leaf stage). After inoculations, plants were kept for 12 hours in mist chambers with repeated misting cycles of 2 minutes on time every 30 minutes and then placed back into growth chambers under the previously described conditions.

### 2.2.3 Analysis of fungal colonization and growth

At 2, 4, and 6 dpi *T. aestivum* and *B. distachyon* leaves were sampled and cut into 1 cm sections before staining with Wheat Germ Agglutinin Alex Fluor® 488 conjugate (WGA-FITC; ThermoFisher Scientific) following previously described procedures (Omidvar et al. 2018). Time points to represent stages of *Pgt* infection were selected based on previous characterization (Figueroa et al. 2013, 2015). To determine the level of fungal colonization, the percentage of urediniospores that germinated (GS), formed an appressorium (AP), established a colony (C), and differentiated a sporulating colony (SC) were visualized using a fluorescence microscope (Leica model DMLB; 450-490 nM excitation). The progression of fungal growth was recorded for 100 infection sites for each of the three biological replicates. Genomic DNA was extracted from *T. aestivum* (three infected primary leaves) and *B. distachyon* (three infected secondary leaves) using the DNeasy Plant Mini Kit (Qiagen) and were standardized to a 10 ng/µl concentration. ITS-specific primers provided

by the Femto™ Fungal DNA Quantification Kit (Zymo Research) were used to amplify ITS regions by qPCR following the manufacturer's instructions; this allowed the quantification of the relative abundance of fungal DNA for the three biological replicates. The *GAPDH* housekeeping gene from each species was used as an internal control to normalize fungal DNA quantities (Omidvar et al. 2018).

### 2.2.4 RNA isolation, purification, and sequencing

Infected and mock treated primary leaves from 'W2691' and 'W2691+*Sr9b*' and secondary leaves from 'Bd21-3' were collected at 2, 4, and 6 dpi. For each of the three biological replicates, three infected leaves were pooled for RNA extraction using the RNeasy Plant Mini Kit (Qiagen). Subsequently, stranded-RNA libraries were constructed, and 125 bp paired-end reads were sequenced on an Illumina HiSeq™ 2500 instrument at the University of Minnesota Genomics Center. On average, more than 10 million reads were generated per time point in each of the previously listed plant-rust interactions (Table S1).

### 2.2.5 Alignment of reads to the *T. aestivum* and *B. distachyon* reference genomes

Short reads and low-quality bases were trimmed using cutadapt v1.18 (Martin 2011) with the following parameters: minimum-length 40, quality-cutoff 30, and quality-base=33. Subsequently, 'W2691' and 'W2691+*Sr9b*' reads were mapped to the *T. aestivum* cv. Chinese Spring reference genome IWGSC RefSeq v1.0 (Alaux et al. 2018) and the 'Bd21-3' reads were mapped to the Bd21-3 reference genome from the Joint Genome Institute (*B. distachyon* Bd21-3 v1.1 DOE-JGI, http://phytozome.jgi.doe.gov/). Read mapping was

conducted using STAR v2.5.3 (Dobin et al. 2012) set for two-pass mapping mode with the following parameters: twopassMode Basic and outSAMmapqUnique 20.

### 2.2.6 Expression profiling and identification of differentially expressed genes

Read counts were mapped to *T. aestivum* and *B. distachyon* genome features using htseq v.0.11.0 (Anders et al. 2015). Normalized read counts and differential expression (DE) analysis were performed with DESeq2 v1.28.1 (Love et al. 2014). Genes with a |log2 fold change| $\geq$ 1.5 and a *p*-value < 0.05 were identified as differentially expressed genes (DEGs).

### 2.2.7 Gene ontology analysis

Gene ontology (GO) terms were used from GOMAP track data for *T. aestivum* (Alaux et al. 2018) and previously published data for *B. distachyon* (*Brachypodium distachyon* Bd21-3 v1.2 DOE-JGI, http://phytozome.jgi.doe.gov/) annotation files. GO terms in wheat and *B. distachyon* were mapped to the GOslim plant subset using owltools --map2slim (https://github.com/owlcollab/owltools). GO enrichment analysis for DEGs was performed using the topGO R package using the "weight01" algorithm and fisher test statistic (Alexa and Rahnenfuhrer 2020). Enriched terms were considered significant when Fisher test p-value < 0.01 (Table S2). Enrichment analyses using the GOslim subset were performed on all differentially expressed wheat and *B. distachyon* genes, as well as on genes within the *S*-gene orthologs clusters. Enrichment analysis with the full GO set was only performed on the differentially expressed *T. aestivum* and *B. distachyon* genes using the same methods described above.

### 2.2.8 Orthology analysis

Protein sequences from *S* genes of interest as cited in original publications (Table S3) were cross-checked using gene name and synonym information and the Basic Local Alignment Search Tool (BLAST) functions in the TAIR gene search database (https://www.arabidopsis.org/index.jsp), EnsemblPlants (https://plants.ensembl.org/index.html), UniPro (https://www.uniprot.org/), and the IPK blast server (https://webblast.ipk-gatersleben.de/barley_ibsc/). OrthoFinder version 2.4.0 (Emms and Kelly 2019) was used to identify orthologs between *A. thaliana* (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Athaliana), *H. vulgare* (http://floresta.eead.csic.es/rsat/data/genomes/Hordeum_vulgare.IBSCv2.36/genome/Hordeum_vulgare.IBSCv2.36.pep.all.fa), *T. aestivum* (annotation version 1.1 https://urgi.versailles.inra.fr/download/iwgsc/IWGSC_RefSeq_Annotations/v1.1/iwgsc_refseqv1.1_genes_2017July06.zip), and *B. distachyon* (annotation version 1.2 https://phytozome-next.jgi.doe.gov/info/BdistachyonBd21_3_v1_2) proteins. For genes in *A. thaliana*, *H. vulgare*, and *T. aestivum* with multiple isoforms, a perl script was used to retain only the longest representative transcript for use in the orthology analysis (https://github.com/henni164/stem_rust_susceptibility/longest_transcript/perl). The longest transcript file for *B. distachyon* (BdistachyonBd21_3_537_v1.2.protein_primaryTranscriptOnly.fa) was obtained from Phytozome. The default settings of OrthoFinder were used, and orthologs of the four species were obtained in a single run. The URGI BLAST tool (https://wheat-urgi.versailles.inra.fr/Seq-Repository/BLAST) was used to identify candidates for missing subgenome representatives.

### 2.2.9 Protein sequence phylogenetic analysis

Using the longest protein sequence from known *S* genes in *A. thaliana* (Lamesch et al. 2011) and *H. vulgare* (Howe et al. 2019), as well as the longest protein sequences from the orthologous candidate *S* genes in *T. aestivum* and *B. distachyon* (Table S3), phylogenetic trees were constructed to examine the relationship of ortholog families using the web-based NGPhylogeny (Lemoine et al. 2019). Default parameters for the FastME one-click workflow were used for MAFFT alignment, BMGE curation, and FASTME tree inference (https://ngphylogeny.fr/documentation). A R script using the packages ggplot2, ggtree, and ape was used to generate visualizations of the generated phylogenetic trees (Wickham 2016; Yu et al. 2017; Paradis and Schliep 2018).

### 2.2.10 Gene co-expression network analysis

Individual gene co-expression networks (GCNs) were constructed and analyzed for *T. aestivum* 'W2691', 'W2691+*Sr9b*' and *B. distachyon* 'Bd21-3' genotypes using the python package Camoco (https://github.com/schae234/Camoco). To build each network, all three independent RNA-seq replicates from all three time points (2, 4, and 6 dpi) of infected and mock-inoculated treatments were used. HTSeq read counts were converted to FPKM values for Camoco compatibility, and then subjected to inverse hyperbolic sine transformation normalized against median FPKMs across all samples. Genes with coefficient of variation < 0.1 across all samples or without a single sample having an expression above 0.5 FPKM were removed from analysis. Additionally, genes with a FPKM value > 0.001 across 60% of samples were included in network analysis. Pearson correlation metrics between all gene pairs were calculated and subjected to Fisher

transformation to generate Z-scores with a cutoff of Z >= 3 to allow comparisons between networks (Huttenhower et al. 2006). Finally, correlation metrics were used to build weighted gene coexpression networks. Clusters containing susceptibility gene orthologs were visualized using ggplot2 (Wickham 2016), ggnetwork (Briatte 2020), sna (Butts 2019), and network (Butts 2015) R packages.

#### 2.2.11 Data availability

Sequence data was deposited in NCBI under BioProject PRJNA483957 (Table S1). Unless specified otherwise, scripts and files for analysis and visualizations are available at https://github.com/henni164/stem_rust_susceptibility. Due to large file sizes, Table S8 and S9 can only be found on the github page mentioned above.

## 2.3 Results

#### 2.3.1 *T. aestivum* and *B. distachyon* differ in susceptibility to *Pgt*

We compared the infection and colonization of *Pgt* in two *T. aestivum* isogenic lines that were susceptible ('W2691') or resistant ('W2691+*Sr9b*') to *Pgt* as well as in the non-host *B. distachyon* Bd21-3. The Bd21-3 line was selected to ease future adoption of reverse genetics approaches, as an extensive collection of T-DNA insertional mutants is available in the same background (Bragg et al. 2012). Symptom development upon infection was consistent with previous observations reporting susceptibility of 'W2691' and 'W2691+*Sr9b*' mediated resistance (intermediate) to race (pathotype) SCCL (Figure 2.1A, B**)** (Zambino et al. 2000). Susceptibility was manifested by formation of large sporulating

pustules in 'W2691', while small pustules surrounded by a chlorotic halo were characteristic of *Sr9b* mediated-resistance at 6 days post-inoculation (dpi). Susceptibility differences between 'W2691' and 'W2691+*Sr9b*' were evident at 6 dpi as formation of fungal colonies was present in both genotypes, but colony sizes were larger in 'W2691' than 'W2691+*Sr9b*' (Figure 2.1). *B. distachyon* supports the formation of colonies that are smaller than those in the resistant *T. aestivum* line 'W2691+*Sr9b*' with no visible macroscopic symptoms observed at 6 dpi (Figure 2.1C). To monitor the progression of fungal growth and colonization, we quantified the percentage of germinated urediniospores (GS), and interaction sites displaying the formation of appressoria (AP), colony formation (C), and colony sporulation (SC) at 2, 4, and 6 dpi using microscopy (Figure 2.1D). The germination frequency (~95%) was similar between all three genotypes tested (ANOVA test, $p > 0.05$). The percentage of interaction sites showing appressorium formation (AP) was higher in wheat than in *B. distachyon* at 4 dpi (ANOVA test, $p \leq 0.035$). The genotype W2691 displayed the highest percentage of interaction sites showing colony formation at 4 and 6 dpi (ANOVA test, $p \leq 0.002$), and sporulation at 6 dpi (ANOVA test, $p \leq 0.0015$). In contrast, a smaller number of rust colonies formed in *B. distachyon*, and these colonies did not show signs of sporulation. To estimate rust colonization levels on *T. aestivum* and *B. distachyon*, we quantified the abundance of fungal DNA in infected leaves at 2, 4, and 6 dpi (Figure 2.1E). Fungal DNA abundance (rust colonization) among all genotypes was not significantly different at 2 dpi (ANOVA test, $p > 0.05$); however, there was a trend at 4 and 6 dpi for higher rust colonization in 'W2691' than in 'W2691+*Sr9b*' and *B. distachyon* (ANOVA test, $p > 0.05$).

### 2.3.2 Putative biological processes associated with *in planta* responses to *Pgt*

The transcriptome profiles of *T. aestivum* ('W2691' and 'W2691+*Sr9b*') and *B. distachyon* (Bd21-3) in response to *Pgt* infection at 2, 4, and 6 dpi were examined using RNA-seq expression profiling (Table S1). Differential expression analysis was used to compare responses to rust infection relative to the baseline mock treatments. Overall, the number of differentially expressed genes (DEGs) increased in 'W2691', 'W2691+*Sr9b*', and 'Bd21-3' over the course of infection (Table 2.1). Between 11-12.9% of *T. aestivum* genes were differentially expressed at 6 dpi, whereas in Bd21-3 only 6.2% were differentially expressed. We conducted a GOslim enrichment analysis on up- and down-regulated DEGs for each interaction at the infection time points (Figure 2.2). At 2 dpi, 'W2691' and 'W2691+*Sr9b*' had few GO terms enriched in either up- or down-regulated DEGs. At 4 dpi, greater similarities between the *T. aestivum* genotypes emerged with very similar enrichment patterns in GOslim terms. The similarity of GO term enrichment continued at 6 dpi, with 'W2691' and 'W2691+*Sr9b*' having nearly identical enrichment patterns. 'W2691+*Sr9b*' had one additional term enriched in both up-regulated (cytoplasm, GO:0005737) and down-regulated (chromatin binding, GO:0003682) genes. Compared to the two *T. aestivum* genotypes, 'Bd21-3' had fewer terms enriched across all three timepoints and only a few terms were in common with 'W2691' and 'W2691+*Sr9b'* (i.e., extracellular region (GO:0005576), DNA-binding transcription factor activity (GO:0003700)). Bd21-3 had several unique terms in both up- and down-regulated categories, among them Mitochondrion (GO:0005739), transporter activity (GO:0005215), catalytic activity (GO:0003824), and DNA binding (GO:0003677) were upregulated, while intracellular (GO:0005622), DNA-binding transcription factor activity (GO:0003700), catalytic activity (GO:0003824), and DNA binding (GO:0003677) were downregulated.

The full GO set also demonstrated clear differences between the *T. aestivum* genotypes and 'Bd21-3'. Photosynthesis-related terms such as chloroplast photosystem I and II (GO:0030093 and GO:0030095), photosystem II antenna complex (GO:0009783), and PSII associated light-harvesting complex II (GO:0009517) were overrepresented at 4 and 6 dpi in 'W2691' and 'W2691+*Sr9b*', but not in 'Bd21-3' (Table S2). In addition, 'Bd21-3' only had enrichment in 11 terms across the cellular component (CC), biological process (BP), and molecular function (MF) categories as compared to the 741 across the three categories in 'W2691' and 'W2691+*Sr9b*' (Table S2). Overall, this analysis highlights how the molecular and genetic responses of 'Bd21-3' to *Pgt* differ from those in 'W2691' and 'W2691+*Sr9b*' over the course of the experiment.

### 2.3.3 Differential regulation of candidate orthologous susceptibility (S) genes in *T. aestivum* and *B. distachyon* upon *Pgt* infection.

Various *S* genes have been previously characterized or postulated in several species, including *A. thaliana* and *H. vulgare* (Büschges et al. 1997; Chen et al. 2007, 2010; Low et al. 2020), and this knowledge has allowed us to further understand molecular plant-microbe interactions. With an interest in identifying potential *S* genes in *T. aestivum* as well as creating resources to enable future studies, we designed a simple experimental workflow based on the identification of known *S* gene orthologs, gene expression comparisons and coexpression network analysis (Figure 2.3). A curated set of previously characterized or postulated *S* genes as summarized by Schie and Takken (2014) was narrowed down by selecting genes in *A. thaliana,* and *H. vulgare*, and eliminating *S* genes that were discovered or characterized for viruses or necrotrophic fungi, leaving 112 potential candidate *S* genes to examine (Table S3). We then conducted an orthology

analysis using all *H. vulgare*, *A. thaliana, B. distachyon,* and *T. aestivum* transcripts to identify orthogroups of longest transcript of all genes. Orthogroups were constructed from 211,973 genes across these species (Table S3). A total of 182,206 genes were assigned to 29,420 orthogroups, the largest of which (OG0000000) contained 211 genes. Of the total genes, 92,913 (86%) wheat, 31,334 *B. distachyon* (80%), 34,075 barley (91%), and 23,883 *A. thaliana* (87%) genes were assigned to orthogroups. We identified 91 of the reported *S* genes from *A. thaliana* and *H. vulgare* across 70 orthogroups, that also consisted of at least one *T. aestivum* gene and one *B. distachyon* gene (Table S4). These genes from *T. aestivum* and *B. distachyon* were selected as *S* gene orthologs. A total of 29,767 genes (orthogroup OG0029421 to OG0059187) were assigned groups with only one member (singleton orthogroups) (Table S5).

The gene expression patterns of *S* gene orthologs in *T. aestivum* and *B. distachyon* were used to identify which may act as susceptibility factors (Figure 2.3, Table S6). The selection criterion was for DEGs that showed a progressive increase in log2 fold change (mock vs infected, $|\log2 \text{ fold change}| \geq 1.5$ and a $p$-value $< 0.05$) in 'W2691' or in both 'W2691' and 'W2691+*Sr9b*', but the corresponding orthologs in *B. distachyon* and/or 'W2691+*Sr9b*' showed a decrease or no change, as observed in various systems (Chen et al. 2010; Pessina et al. 2014). The assumption is that *S* genes will be up-regulated during infection when the pathogen reaches the sporulation stage (e.g., in a susceptible or intermediate resistant host represented by 'W2691' and 'W2691+*Sr9b*', respectively), but with a low or no regulatory change in a non-host (Bd21-3). Expression data for all genes can be found in Table S6 in association with orthogroup number. Most genes in the 70 orthogroups did not demonstrate major changes in expression over the course of the

experiment (Appendix F), including the orthogroup OG0001703, which contains the *Mlo* alleles and orthologous sequences. Eight orthogroups that demonstrated these expression patterns were chosen for further analysis; these included ortholog genes for *AGD2* (*aberrant growth and death 2*), *BI-1* (*BAX inhibitor-1*), *DMR6* (*downy mildew resistance 6*), *DND1* (*defense*, *no death*), *FAH1* (*fatty acid hydroxylase 1*), *IBR3* (*IBA response 3*), *VAD1* (*vascular associated death 1*), and *WRKY25* (*WRKY DNA binding protein 25*) (Figure 2.4, Table 2.2, Table S7). Among the eight susceptibility orthogroups, *T. aestivum* orthologs of *BI-1*, *DMR6,* and *WRKY25* showed the greatest increase in fold change (Table S7) in either 'W2691' or 'W2991+*Sr9b*', particularly at 6 dpi (Figure 2.4). The gene ortholog of *DND1* displayed a higher fold change in 'W2691' than in 'W2691+*Sr9b*'.

The phylogenetic relationships of the orthogroups to known *S* genes were confirmed using NGphylogeny (Appendix F). A phylogenetic tree for *DND1* was not generated since the orthogroup (OG0018857) only contains three genes (TraesCS5D01G404600, BdiBd21-3.1G0110600, and AT5G15410). Complete sets of *T. aestivum* homeologs from the three subgenomes were found in four out of the eight examined orthogroups. There are only two of three expected *T. aestivum* homeologs in the *DMR6* orthogroup, with TraesCS4B02G346900 and TraesCS4D02G341800 representing the B and D subgenomes, respectively. A tblastn of these sequences to chromosome 4A revealed TraesCS4A02G319100, a partial match of 30-31% identity (1e-42 to 1e-44). This gene has low expression and is found in orthogroup OG0006808, which contains two other *T. aestivum* genes, one *B. distachyon* gene, two *A. thaliana* genes, and one *H. vulgare* gene (Tables S4 and S6). Despite the low sequence similarity, TraesCS4A02G319100 and

TraesCS4B02G346900 are at more similar positions (4A:608043459 and 4B:640532917, respectively) to each other than to TraesCS4D02G341800 (4D:498572979). A tblastn to the entire genome revealed 20 other matches for the two *DMR6* orthologs with 31-74% identity. Thus, it does not seem that the *T. aestivum* genome reference (Chinese Spring) contains a homeolog of *DMR6* in the A genome. To further examine if a *DMR6* ortholog is present in the A subgenome, the B and D subgenome *DMR6* homeologs were BLASTed to the wheat pangenome CDS sequences (https://galaxy-web.ipk-gatersleben.de/). Similar results were obtained; there were 30 total hits ranging from 36.765 to 39.564% identity (1.07e-61 to 1.99e-69) on chromosome 4A of the 10 genomes. Based on the low sequence identity, it is likely that there is not an ortholog of *DMR6* on chromosome 4A. However, hits with high identity (97.619-98.214, e = 0) were found on chromosome 5A in all 10 genomes.

Another *S* gene orthogroup without full subgenome representation was OG0018857 which represents *DND1*. This orthogroup only has one *T. aestivum* gene, TraesCS5D02G404600 from subgenome D. A tblastn to chromosomes 5A and 5B resulted in matches with high identity on both 5A (TraesCS5A02G395300, 94%, 6e-159) and 5B (89%, 1e-176). TraesCS5A02G395300 is present in orthogroup OG0048986 as a singleton with low expression in 'W2691' and 'W2691+*Sr9b*', and TraesCS5B02G400100 is included in orthogroup OG0048844 as a singleton as well with some notable expression at 6 dpi in infected 'W2691' and low expression in 'W2691+*Sr9b*' (Table S5 and S6). A tblastn to the entire genome identified 19 other candidates with identity 33-97%. The most notable matches with high identity are TraesCS7B02G161600 (97%, 4e-152) and

TraesCS3B02G306700 (97%, 2e-147), which are the only two genes together in orthogroup OG0027858. Both top matches had essentially no expression in either *T. aestivum* genotype (Table S6). A third genomic region on chromosome 2B also has 97% identity, but it is annotated as a nested repeat. For *FAH1*, three *T. aestivum* orthologs are present in the orthogroup OG0006155, but one is from subgenome A (TraesCS5A02G019200) while the other two are from subgenome D (TraesCS5D02G024600 and TraesCS5D02G424200). A tblastn of all three sequences to chromosome 5B revealed two matches, TraesCS5B02G016700 (87%, 3e-49) and TraesCS5B02G418800 (62/87%, 2e-64/2e-49), while a tblastn to the entire genome uncovered a partial match on 5A (TraesCS5A02G416500, 46-62%, 7e-68-1e-104) and a partial match on 3D which was not annotated (68-89%, 6e-34-3e-43). All three annotated genes are in singleton orthogroups (TraesCS5B02G016700, OG0056228; TraesCS5B02G418800, OG0055841; TraesCS5A02G416500, OG0057903) and have low expression in both 'W2691' and 'W2691+*Sr9b*' (Table S5 and S6). Orthogroup OG0005265 for *AGD2* is similar to the orthogroup for *FAH1*, having one A subgenome representative (TraesCS4A02G116000) and two D subgenome representatives (TraesCS4D02G189600 and TraesCS7D02G452900). The tblastn of these sequences to chromosome 4B revealed one possible match with two annotations in the same location (61-62%, 9e-108-1e-113), TraesCS4B02G264500 on the - strand and TraesCS4B02G264400 on the + strand. The former is a singleton in orthogroup OG0047603 with low expression in 'W2691' and high expression in infected 'W2691+*Sr9b*' at 6 dpi, while the latter is in OG0015484 with several other genes and is

not highly expressed in either *T. aestivum* genotype (Table S6). The tblastn to the entire genome revealed several hits of identity varying between 22% and 97%.

### 2.3.4 Gene coexpression network analysis

To further explore potential processes and novel genes linked to stem rust susceptibility, a gene coexpression network for *B. distachyon* and each *T. aestivum* genotype using the mock and infected RNA-seq data at each timepoint was constructed (Table S8, see github page). The complete 'Bd21-3' network had 572,179 edges that connected 21,746 nodes (55.7% of protein-coding genes), while the 'W2691' and 'W2691+*Sr9b*' networks were larger (W2691: 3,433,279 edges, 49,082 nodes, 45.6% of protein-coding genes; W2691+*Sr9b*: 3,817,404 edges, 49,000 nodes, 45.5% of protein-coding genes). The *B. distachyon* network was expected to be smaller as it represents a diploid species with fewer annotated genes (39,068), while the hexaploid wheat contains more gene annotations (107,891). There were 189 clusters with more than 10 genes in 'Bd21-3', 258 in 'W2691', and 391 in 'W2691+*Sr9b*'. Thus, more genes had similar expression patterns in 'W2691+*Sr9b*' than in 'W2691', and 'Bd21-3' had the lowest number of genes with similar patterns. The eight *S* gene orthogroups of interest were represented by 14 clusters in 'W2691' (cluster IDs: 0, 3, 4, 5, 8, 11, 13, 60, 110, 178, 11114, 11235, 12377, 20128), 11 in 'W2691+*Sr9b*' (cluster IDs: 0, 2, 4, 112, 1139, 1916, 2729, 2772, 3133, 10229, 11079), and 11 in Bd21-3 (cluster IDs: 3, 4, 35, 51, 272, 513, 652, 1359, 1662, 1848, 3087) (Table S9, see github page). Some orthogroups were represented across multiple clusters, while others were only represented in singleton clusters. The ortholog clusters in *B. distachyon* contained fewer genes than the corresponding 'W2691' and 'W2691+*Sr9b*'ortholog clusters.

GO enrichment tests using GOslim annotations were conducted on the clusters to investigate functional processes. Across all eight *S* gene orthogroups, at least one gene from each was in a cluster with GO enrichment in at least one genotype (Figure 2.5). *DMR6*, *FAH1*, and *WRKY25* were the only candidate(s) to have enrichment in all three genotypes, *AGD2* and *DND1* only had enrichment in 'W2691', and *BI-1*, *IBR3*, and *VAD1* had enrichment in both 'W2691' and 'W2691+*Sr9b'*. Terms commonly enriched in the *T. aestivum* genotypes include the Golgi apparatus (GO:0005794), endosome (GO:0005768), endoplasmic reticulum (GO:0005783), protein binding (GO:0005515), transporter activity (GO:0005215), vacuole (GO:0005773), and peroxisome (GO:0005777) (Figure 2.5). Only one GO term, catalytic activity (GO:0003824) was unique to 'Bd21-3', with other terms like DNA-binding transcription factor activity (GO:0003700) being enriched in Bd21-3 and *T. aestivum* genotypes.

For each genotype, a cluster containing one or more orthologs of *DND1*, *VAD1*, and *DMR6* was selected as examples for presentation (Figure 2.6). Selection criteria for these examples included 1) higher expression in infected than in mock treatments in *T. aestivum* and 2) varied cluster sizes across genotypes. *DND1* was represented by TraesCS5D02G404600 within cluster 4 in the 'W2691' genotype (557 genes), by TraesCS5D02G404600 within cluster 122 in the 'W2691+*Sr9b*' genotype (21 genes) and by BdiBd21-3.1G0110600 within cluster 652 in the Bd21-3 genotype (4 genes) (Figure 2.6A, Table S8). *VAD1* represented a mid-point between *DND1* and *DMR6*, with the large cluster 0 (TraesCS2D02G236800) representing VAD1 for the 'W2691' genotype (4527 genes), a singleton cluster (cluster 3087) for the Bd21-3 genotype (1 gene, BdiBd21-3.1G0357000), and the large cluster 0 (TraesCS2D02G236800) for the 'W2691+*Sr9b*'

genotype (3400 genes) (Figure 2.6B, Table S8). *DMR6* was also represented by cluster 0 (TraesCS4B02G346900 and TraesCS4D02G341800) for both 'W2691' and 'W2691+*Sr9b*'; however, cluster 4 representing *DMR6* in 'Bd21-3' (BdiBd21-3.1G1026800) was larger than in the previous examples (443 genes) (Figure 2.6C, Table S8). In all cases, the *S* gene candidates were not the most differentially-expressed genes at 6 dpi among the *T. aestivum* genotype clusters; the most differentially expressed gene at 6 dpi in cluster 0 was TraesCS7A02G157400 (not functionally annotated) in 'W2691' and TraesCS1A02G266000 (IPR002921:Fungal lipase-like domain IPR029058:Alpha/Beta hydrolase fold IPR033556:Phospholipase A1-II) in W2691+*Sr9b*. For cluster 4 in W2691, TraesCS4D02G120200 (IPR001471:AP2/ERF domain IPR016177:DNA-binding domain superfamily IPR036955:AP2/ERF domain superfamily) is the most differentially expressed gene at 6 dpi, while for cluster 122 it is TraesCS1A02G276800 (IPR013087:Zinc finger C2H2-type IPR036236:Zinc finger C2H2 superfamily). In *B. distachyon*, BdiBd21-3.1G0110600, which is the Bd21-3 ortholog to *A. thaliana DND1*, was most differentially expressed in the cluster representing *DND1* and was highly downregulated in infected tissue at 6 dpi. By necessity the most differentially expressed gene in the network representing *VAD1* in *B. distachyon* is the ortholog of *VAD1*, as Bd21-3 cluster 3087 is a singleton cluster. The most differentially expressed Bd21-3 gene in cluster 4 representing *DMR6* is BdiBd21-3.2G0466100. This gene is annotated as a leucine-rich repeat protein kinase family protein due to homology with the *A. thaliana* gene AT1G79620, though the orthology analysis places these genes in different clusters (OG0019394 and OG0010938, respectively). All clusters representing the eight *S* gene candidates are shown in Appendix F.

## 2.4 Discussion

Susceptibility (*S*) genes are an essential component of compatible plant pathogen interactions (Engelhardt et al. 2018). The opportunity to genetically manipulate such genes to engineer disease resistance in important crops such as *T. aestivum* has captured significant scientific interest in recent years. However, our understanding of the genetic basis of disease susceptibility in cereals is limited to a few examples (van Schie and Takken 2014; Engelhardt et al. 2018). Thus, important questions regarding the biological functions of these genes and their activation remain to be answered. As a first step to uncover putative stem rust *S* genes, we conducted a comparative RNA-seq experiment coupled with gene co-expression network analysis to determine transcriptional responses in *T. aestivum* genotypes and *B. distachyon* Bd21-3. We compared a compatible interaction (W2691) with an incompatible interaction controlled by the race-specific resistance gene *Sr9b* in the same genetic background (W2691+*Sr9b*). *Sr9b* restricts pathogen growth; however, it also allows the development of small sporulating colonies of a *Pgt* isolate which belongs to the race SCCL (Zambino et al. 2000). A more stringent incompatibility scenario is given by Bd21-3 genotype of *B. distachyon*, which allows restricted colony formation of *Pgt* without sporulation. These observations were consistent with previous descriptions of *B. distachyon* as a non-host to rust pathogens (Figueroa et al. 2013, 2015; Omidvar et al. 2018). Thus, a strength of this study is that we surveyed molecular responses associated with increasing levels of susceptibility.

Consistent with findings from other transcriptomic studies of wheat-rust interactions (Manickavelu et al. 2010; Zhang et al. 2014; Dobon et al. 2016; Chandra et al. 2016; Yadav

et al. 2016), major transcriptional changes were detected in response to infection in both *T. aestivum* and *B. distachyon*, which reflect the complexity of these plant-microbe interactions. A significantly higher number of up- or down-regulated genes were found in *T. aestivum* than *B. distachyon*. The greater fungal colonization of *T. aestivum* as indicated by *in planta* fungal growth assays of *Pgt* is likely a result of the pathogen's failure to effectively manipulate the metabolism of *B. distachyon*. GOslim term analyses indicated an enrichment for Golgi apparatus, peroxisome, vacuole, and cell wall related functions in up-regulated genes in *T. aestivum*. These results are not surprising as a large proportion of immune receptors and plant defense signaling components play a role in plant-microbe interactions (Dodds and Rathjen 2010; Couto and Zipfel 2016). The plant Golgi apparatus and peroxisomes have been reported as targets of effectors from various pathogenic filamentous fungi (Robin et al. 2018). The enrichment of these GO terms in up-regulated genes in *T. aestivum* suggests that these cellular components may be direct or indirect targets for effectors derived from *Pgt*. Analyses with the full GO term set revealed many enriched terms among downregulated genes related to photosynthesis in 'W2691' and 'W2691+*Sr9b*'; a decrease in chlorophyll and photosynthetic activity has been previously reported in wheat infected with *Pgt* (Berghaus and Reisener 1984; Moerschbacher et al. 1994).

Several *S* genes to diverse pathogens have been identified or postulated in various plant species (van Schie and Takken 2014; Engelhardt et al. 2018). While this area of research for cereal rust pathogens is still in its infancy, positive results from other pathosystems make a strong case to consider the modification of *S* genes as an approach to deliver durable

and broad-spectrum disease resistance. So far, only a few host-delivered avirulence effectors, *AvrSr50,* (Chen et al. 2017), *AvrSr35* (Salcedo et al. 2017), *AvrSr27* (Upadhyaya et al. in press) from any cereal rust fungi have been isolated. These were identified in *Pgt* and how these effectors disrupt defense responses in compatible interactions remains unknown. Future research seeking to identify which plant proteins these effectors target will help elucidating *S* genes or processes required for stem rust susceptibility.

Here, expression patterns of gene orthologs in *T. aestivum* and *B. distachyon* corresponding to previously characterized *S* genes in *H. vulgare,* and *A. thaliana* were examined to develop a framework to study *S* genes in wheat. A key focus of this study was to develop a simple workflow to extract orthologs with high expression in stem rust susceptible *T. aestivum* but low expression in either *T. aestivum* with intermediate resistance, or *B. distachyon.* We note that differential gene expression between *T. aestivum* genotypes (W2691 and W2691+*Sr9b*) can also provide an opportunity to discover *S* genes since rust infections in both genotypes differ. To link these candidate *S* genes with the biological pathways in wheat and *B. distachyon*, we constructed gene coexpression networks which can be further explored to determine the role of components of these pathways and the complex interplay towards regulation of susceptibility in *Pgt*-*T. aestivum* interactions.

The biological functions of *S* genes in compatible-plant microbe interactions are diverse as these genes play roles in a wide array of events that are critical for pathogen accommodation and survival (Engelhardt et al. 2018). Some of these susceptibility genes can act as negative regulators of immune responses, such as PTI, cell death, and

phytohormone-related defense. Our study determined that *T. aestivum* orthologs of the BAX inhibitor-1 (*BI-1*) gene in *H. vulgare*, are candidate *S* genes as these were upregulated in 'W2691' (6 dpi) and 'W2691+*Sr9b*' (4-6 dpi), whereas their expression in *B. distachyon* was not affected. *BI-1* is an endoplasmic reticulum membrane-localized cell death suppressor in *A. thaliana*, and its wheat ortholog *TaBI-1* (accession GR305011) is proposed to contribute to susceptibility in *T. aestivum* to the biotrophic pathogen *Puccinia striiformis* f. sp. *tritici* (Wang et al. 2012). Interestingly, the highest upregulation of the *BI-1* was detected in the 'W2691+*Sr9b*' genotype where it is necessary to regulate a HR upon *Pgt* recognition. Given this result it should be examined if *BI-1* may be a conserved plant *S* factor to wheat rust fungi. Various orthologs of *FAH1*, which encodes a ferulate *5*-hydroxylase in *A. thaliana*, were upregulated in the *T. aestivum* genotypes upon *Pgt* infection (Mitchell and Martin 1997). According to studies in *A. thaliana*, FAH1 plays a role in *BI-1*-mediated cell death suppression through interaction with cytochrome $b_5$ and biosynthesis of very-long-chain fatty acids (Nagano et al. 2012). Additional findings further suggest that *Pgt* can also interfere with cell death signaling by altering *VAD1* expression. The *VAD1* gene encodes a putative membrane-associated protein with lipid binding properties, and it is proposed to act as negative regulator of cell death (Lorrain et al. 2004; Khafif et al. 2017). Transcriptional activation of *VAD1* has been shown to occur in advanced stages in plant pathogen interactions (Bouchez et al. 2007). We detected an upregulation of *VAD1* orthologs in *T. aestivum* at 6 dpi, which is considered a late infection stage in the establishment of rust colonies.

Salicylic acid (SA) is a key phytohormone required to orchestrate responses to many pathogens (Ding and Ding 2020). Similar to VAD1 whose function as a *S* factor is SA-dependent, we also uncovered other upregulated genes that may also participate in defense suppression. The orthologs of the *DMR6* are highly upregulated in *T. aestivum* at 4 and 6 dpi in both compatible and incompatible interactions. As characterized in *A. thaliana, DMR6* encodes a putative 2OG-Fe(II) oxygenase that is defense-associated and required for susceptibility to downy mildew through regulation of the SA pathway (Van Damme et al. 2008; Zhang et al. 2017). The role of *DMR6* in disease susceptibility holds significant promise to control diverse pathogens. For instance, mutations in *DMR6* confer resistance to hemibiotrophic pathogens *Pseudomonas syringae* and *Phytophthora capsici* (Zeilmaker et al. 2015) and silencing of *DMR6* in potato increases resistance to the potato blight causal agent, *P. infestans* (Sun et al. 2016). It has also been shown that the *H. vulgare* ortholog genetically complements *DMR6* knock-out *A. thaliana* lines and restores susceptibility to *Fusarium graminearum* (Low et al. 2020). Gene orthologs of *DND1* were also identified as upregulated in both *T. aestivum* genotypes. The gene *DND1* encodes a cyclic nucleotide-gated ion channel and its activity is also related to SA regulation (Clough et al. 2000). Mutations in *A. thaliana DND1* display enhanced resistance to viruses, bacteria and fungal pathogens (Yu et al. 2000; Jurkowski et al. 2004; Genger et al. 2008; Sun et al. 2017). We also noted that several wheat orthologs of the *A. thaliana* gene *IBR3* also increased in expression as *Pgt* infection advanced. The role of *IBR3* in susceptibility to *P. syringae* in *A. thaliana* has been confirmed by mutations and overexpression approaches (Huang et al. 2013). Consistent with our results, *IBR3* is upregulated in *A. thaliana* upon infection by *P. syringae*.

Plant transcriptional reprogramming triggered by pathogen perception is often mediated by WRKY transcription factors through activation of the MAP kinase pathways (Eulgem and Somssich 2007; Rushton et al. 2010). Here, we detected an upregulation of the expression of *WRKY25* orthologs that was most prominent at 6 dpi in the 'W2691+*Sr9b*' genotype. The Arabidopsis gene *AtWRKY25* is induced in response to the bacterial pathogen *Pseudomonas syringae* and the SA-dependent activity of *AtWRKY25* is also linked to defense suppression (Zheng et al. 2007). According to results from this study, the contribution of orthologs of *AGD2* to stem rust susceptibility in *T. aestivum* should also be examined. *AGD2* encodes an aminotransferase and participates in lysine biosynthesis at the chloroplast (Song et al. 2004). Given that several oomycete and fungal effectors target the chloroplast (Kretschmer et al. 2020), effector research in cereal rust pathogens will be crucial to determine if these pathogens also target this organelle.

A classic example of *S* genes in barley is given by the *Mlo* (*Mildew Locus O*) gene (Jørgensen 1992) in which a recessive mutation results in broad spectrum resistance to *Blumeria graminis* f. sp. *hordei*, the causal agent of powdery mildew. The *Mlo* gene family is highly conserved across monocot and dicot plants and gene editing of *Mlo* homeologs in wheat confers resistance to powdery mildew (Acevedo-Garcia et al. 2017). Interestingly, *Mlo* genes in *T. aestivum* have not been reported to provide protection against cereal rust diseases. Consistent with this, our study did not detect a significant change in the expression of *Mlo* alleles in *T. aestivum* genotypes ('W2691' and 'W2691+*Sr9b*') over the course of the experiment.

One caveat of this study is that some *S* genes in *T. aestivum* for *Pgt* may not be found in model species like *A. thaliana* or detected using other pathogens. However, this is a first step to identify candidates to guide functional studies. While in this study we focused on orthologous *S* genes, the gene co-expression networks presented here are excellent resources to identify additional candidate *S* factors. It is possible that some of the genes included in clusters of these networks are part of the regulatory process that control expression of *S* genes or are part of essential pathways although their function may not be characterized yet in other systems. Future functional studies are required to validate the function of these genes in *T. aestivum* as *S* factors for rust infection and determine if these can be exploited for agricultural practice. A key aspect for the success of these novel approaches is the absence of plant developmental defects resulting from mutations of *S* genes. In some cases, the loss-of-function of negative regulators leads to constitutive activation of plant defense responses that manifest as poor growth or lesion-mimic phenotypes among other pleiotropic effects (Büschges et al. 1997; Ge et al. 2016). VIGS-mediated transient gene silencing (Lee et al. 2015), RNAi-mediated silencing (Helliwell and Waterhouse 2003; Waterhouse and Helliwell 2003; Sun et al. 2016), and TILLING populations include some of the approaches to explore the potential use of these *S* gene candidates. Gene editing technologies through Zinc Finger nucleases, TALENs, CRISPR/Cas9 systems also offer options to generate transgene free plants (Urnov et al. 2010; Gaj et al. 2013; Wang et al. 2014; Jia et al. 2017; Nekrasov et al. 2017; Kim et al. 2018; Luo et al. 2019). In conclusion, as the demands for multi-pathogen durable disease

resistance rise, our ability to target *S* genes may serve as a sound approach to harness genetic diversity and maximize the resources to meet critical these grand challenges.

## 2.5 Data Availability Statement

Sequence data was deposited in NCBI under BioProject PRJNA483957 (Table S1). Unless specified otherwise, scripts and files for analysis and visualizations are available at https://github.com/henni164/stem_rust_susceptibility. Supplemental tables are available at https://github.com/henni164/stem_rust_susceptibility.

**Table 2.1. Differentially expressed genes in *T. aestivum* and *B. distachyon* in response to *P. graminis* f. sp. *tritici* infection.**

| | 2 dpi | | 4 dpi | | 6 dpi | |
|---|---|---|---|---|---|---|
| **Genotype** | **up** | **down** | **up** | **down** | **up** | **down** |
| W2691 | 278 | 577 | 2,887 | 2,614 | 6,659 | 7,241 |
| W2691+*Sr9b* | 747 | 110 | 3,835 | 1,832 | 6,397 | 5,471 |
| Bd21-3 | 200 | 437 | 559 | 1,419 | 739 | 1,665 |

* Total number of wheat genes: 107,891; total number of *B. distachyon* genes: 39,068. Within-genotype comparisons used mock treatments as the baseline.

**Table 2.2. List of *S* genes explored through the gene expression analysis.**

| Gene | Annotation *** | Postulated Mechanism of Susceptibility *** | Pathogen species and Disease*** | Reference |
|---|---|---|---|---|
| AGD2 (AT4G33680*) | Aberrant growth and death 2 | Defense suppression (possibly SA-dependent) | *Pseudomonas syringae* (bacterial speck) | Rate and Greenberg (2001); Song et al. (2004) |
| BI-1 (HORVU6Hr1G014450**) | Bax inhibitor-1 | Membrane rearrangement, | *Blumeria graminis f. sp. hordei* | Eichmann et al. (2010) |

| | | | | |
|---|---|---|---|---|
| | | haustorium establishment, and suppression of cell death | (powdery mildew) | |
| DMR6 (AT5G24530*) | 2-oxoglutarate (2OG)-Fe(II) oxygenase | Defense suppression (SA dependent) | *Hyaloperonospora parasitica* (downy mildew) | van Damme et al. (2005, 2008) |
| DND1 (AT5G15410*) | CNGC2/4 cyclic nucleotide gated channel | Defense suppression and possible regulator of nitric oxide synthesis (SA-dependent) | *Hyaloperonospora parasitica* (downy mildew), *Alternaria brassicicola* (black leaf spot), *Botrytis cinerea* (grey mold/rot), *Pectobacterium carotovorum* (bacterial soft rot), *Pseudomonas syringae* (Bacterial speck) | Govrin and Levine (2000); Ahn (2007); Genger et al. (2008); Su'udi et al. (2011); |
| FAH1 (AT2G34770*) | Fatty acid hydroxylase 1 | Defense suppression (SA dependent) | *Golovinomyces cichoracearum* (powdery mildew) | Konig et al. (2012) |
| IBR3 (AT3G0*6810) | IBA response 3 | Defense suppression PTI (auxin independent) | *Pseudomonas syringae* (bacterial speck) | Huang et al. (2013) |
| VAD1 (AT1G02120*) | Vascular Associated death1 | Defense suppression (SA and ET dependent) | *Pseudomonas syringae* (bacterial speck) | Lorrain et al. (2004); Bouchez et al. (2007) |
| WRKY25 (AT2G30250*) | WRKY DNA-binding protein 25 | Defense suppression (SA dependent) | *Pseudomonas syringae* (bacterial speck) | Zheng et al. (2007) |

Sources:

* TAIR database

** Ensembl Plants

*** from van Schie and Takken, 2014

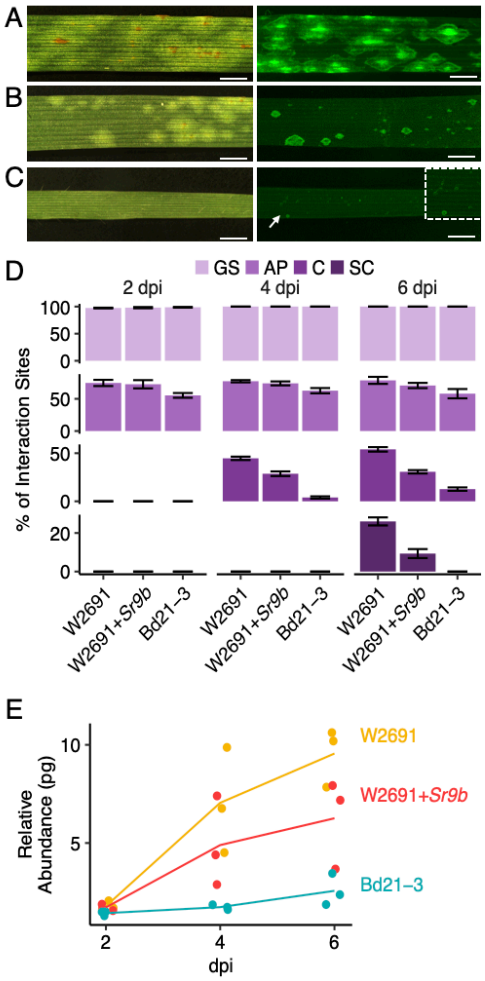



**Figure 2.1.** Infection of *T. aestivum* and *B. distachyon* genotypes with *P. graminis* f. sp. *tritici* race SCCL. (**A-C**) Development of disease symptoms (left) and fungal colonization (right) at 6 dpi. (**A**) W2691 (susceptible wheat line). (**B**) W2691+*Sr9b* (resistant wheat line). (**C**) *B. distachyon* Bd21-3 line (non-host). The white arrow and the white box indicate the area which was enlarged for better visualization of colonies. Scale bars indicate 2 mm. (**D**) Percentage of fungal infection sites which showed germinated urediniospores (GS), appressorium formation (AP), colony establishment (C), and sporulating colony (SC). Error bars represent the standard error of three independent biological replicates. (**E**) Fungal DNA abundance in infected 'W2691', 'W2691+*Sr9b*', and 'Bd21-3' genotypes as

measured using qPCR. Points represent actual sample values and the line represents the mean of the samples.

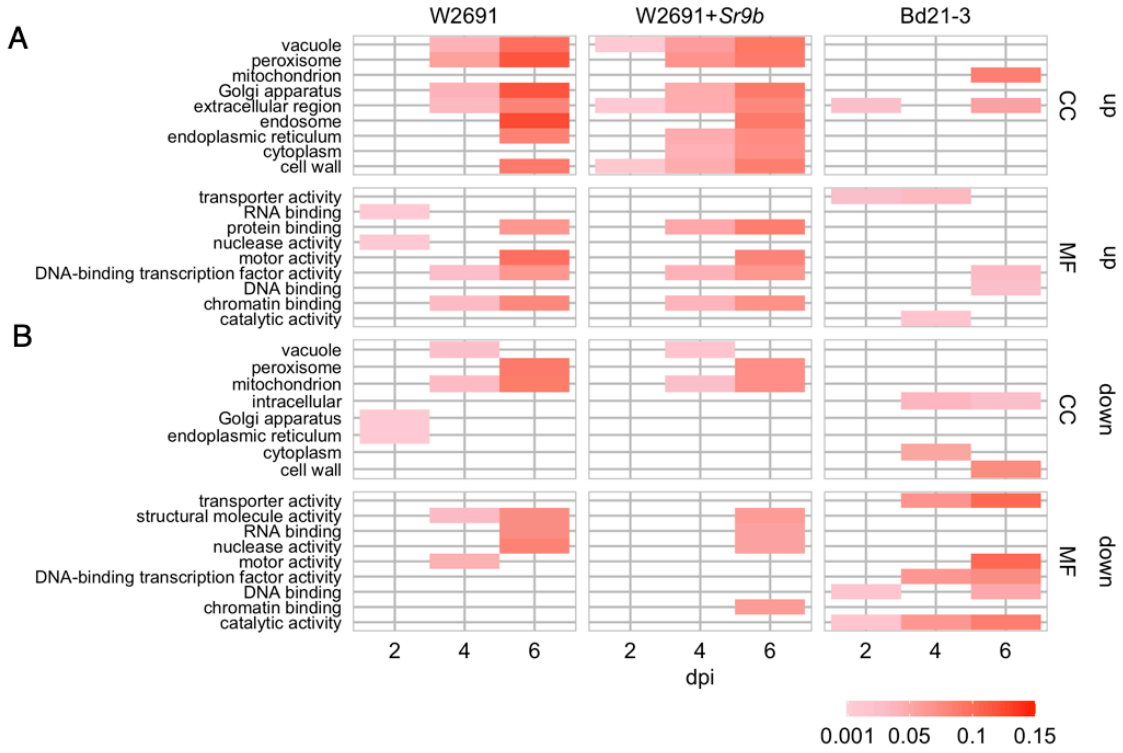



**Figure 2.2.** GOSlim enrichment analysis of differentially expressed (DE) genes in mock vs inoculated *T. aestivum* ('W2691' and 'W2691+*Sr9b*') and *B. distachyon* (Bd21-3) genotypes across three time points (bottom x-axis) upon infection with *P. graminis* f. sp. *tritici*. (**A**) Enrichment of plant GOSlim terms of upregulated (up) DE genes and (**B**) downregulated (down) DE genes. The y-axis shows plant GO slim terms separated by category: cellular component (CC) and molecular function (MF). The scale represents the proportion of genes annotated with each GO term to all the genes tested.

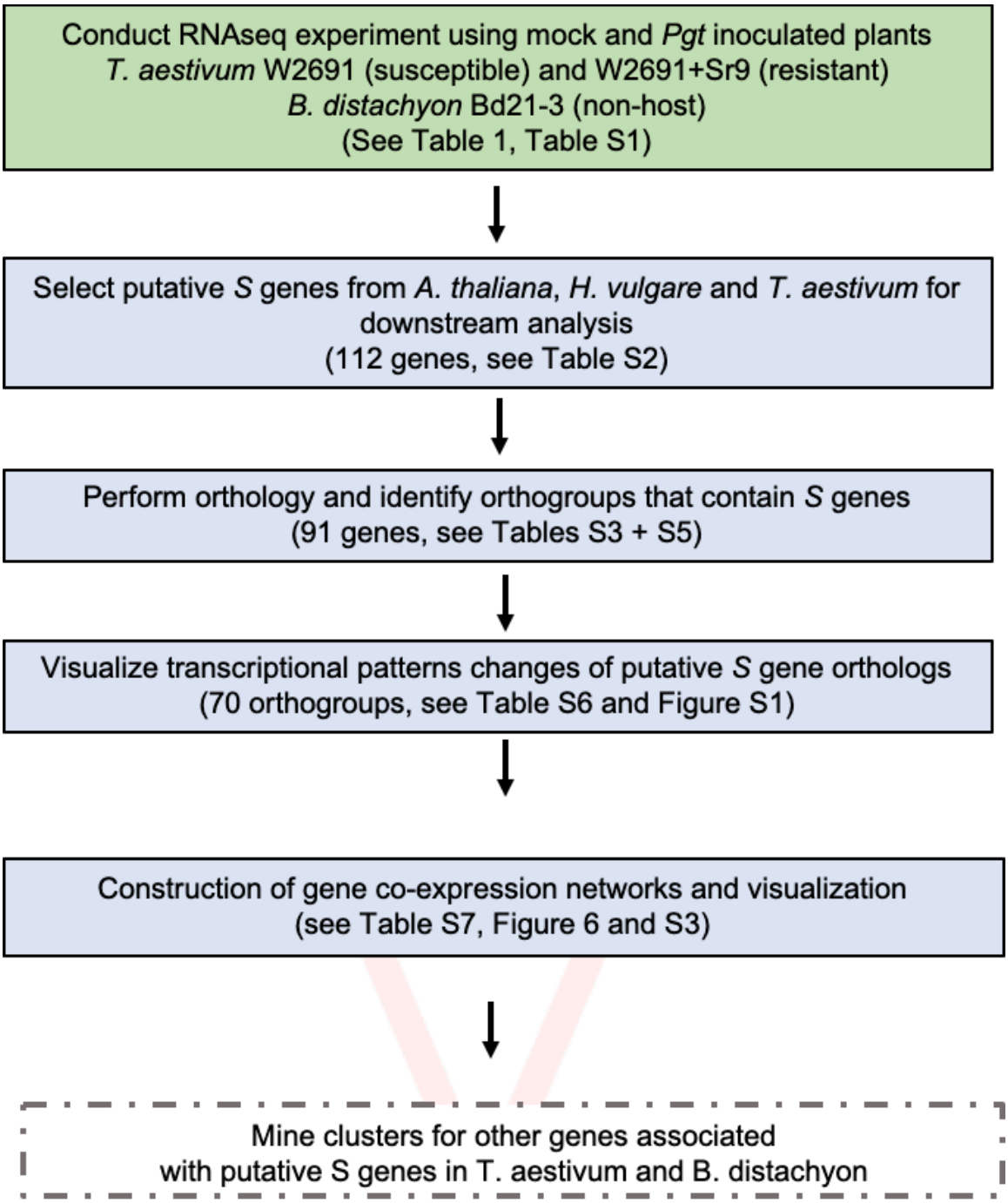



**Figure 2.3.** Experimental workflow used to identify candidates of *S* genes that contribute to infection of *T. aestivum* by *P. graminis* f. sp. *tritici.*

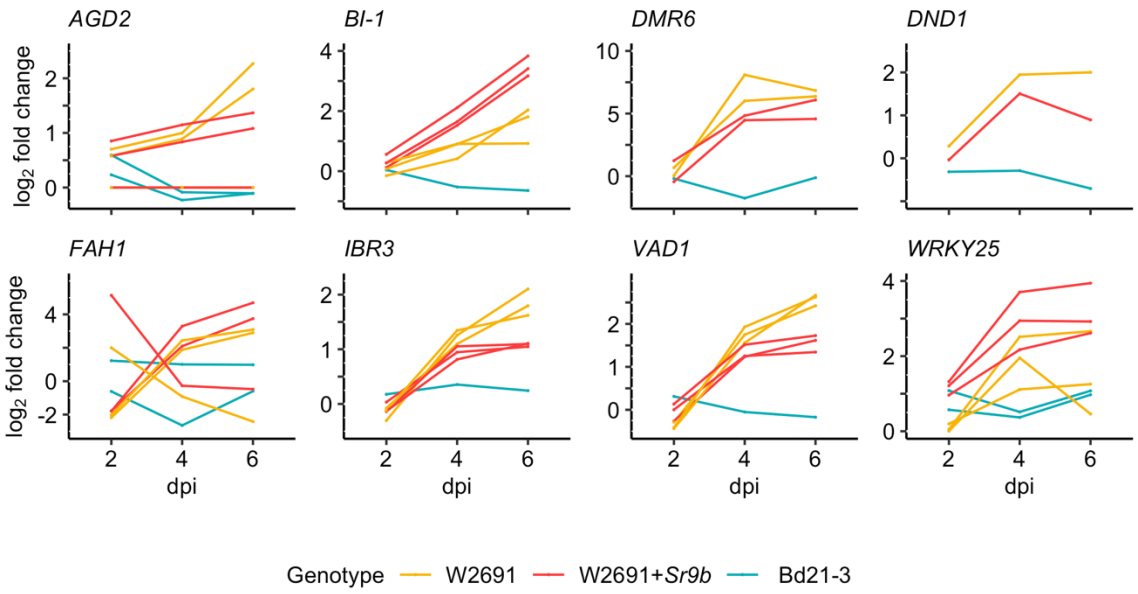



**Figure 2.4.** RNAseq expression profile patterns of selected orthogroups containing candidate *S* genes in *T. aestivum* ('W2691' and 'W2691+*Sr9b*') and *B. distachyon* (Bd21-3) genotypes throughout infection with *P. graminis* f. sp. *tritici*. $Log_2$ fold change values for all gene orthologs are presented for each infected genotype compared to the mock treatment per sampling time point. Gene IDs, average FPKM values, orthogroup, and coexpression cluster identifiers are presented in Table S7.

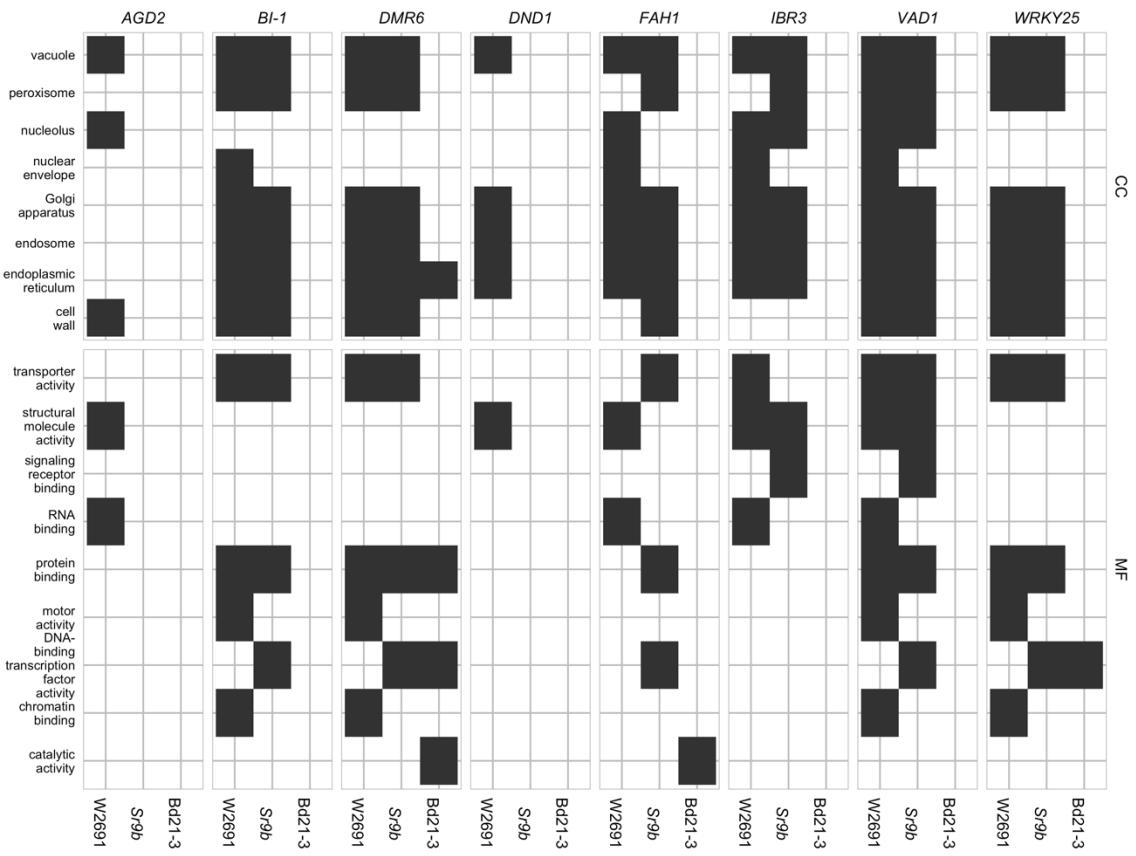



**Figure 2.5.** GO term enrichment for all genes in co-expression gene clusters containing *S* gene orthologs in *T. aestivum* and *B. distachyon*. The y-axis shows GO slim terms separated into categories: cellular component (CC) and molecular function (MF).

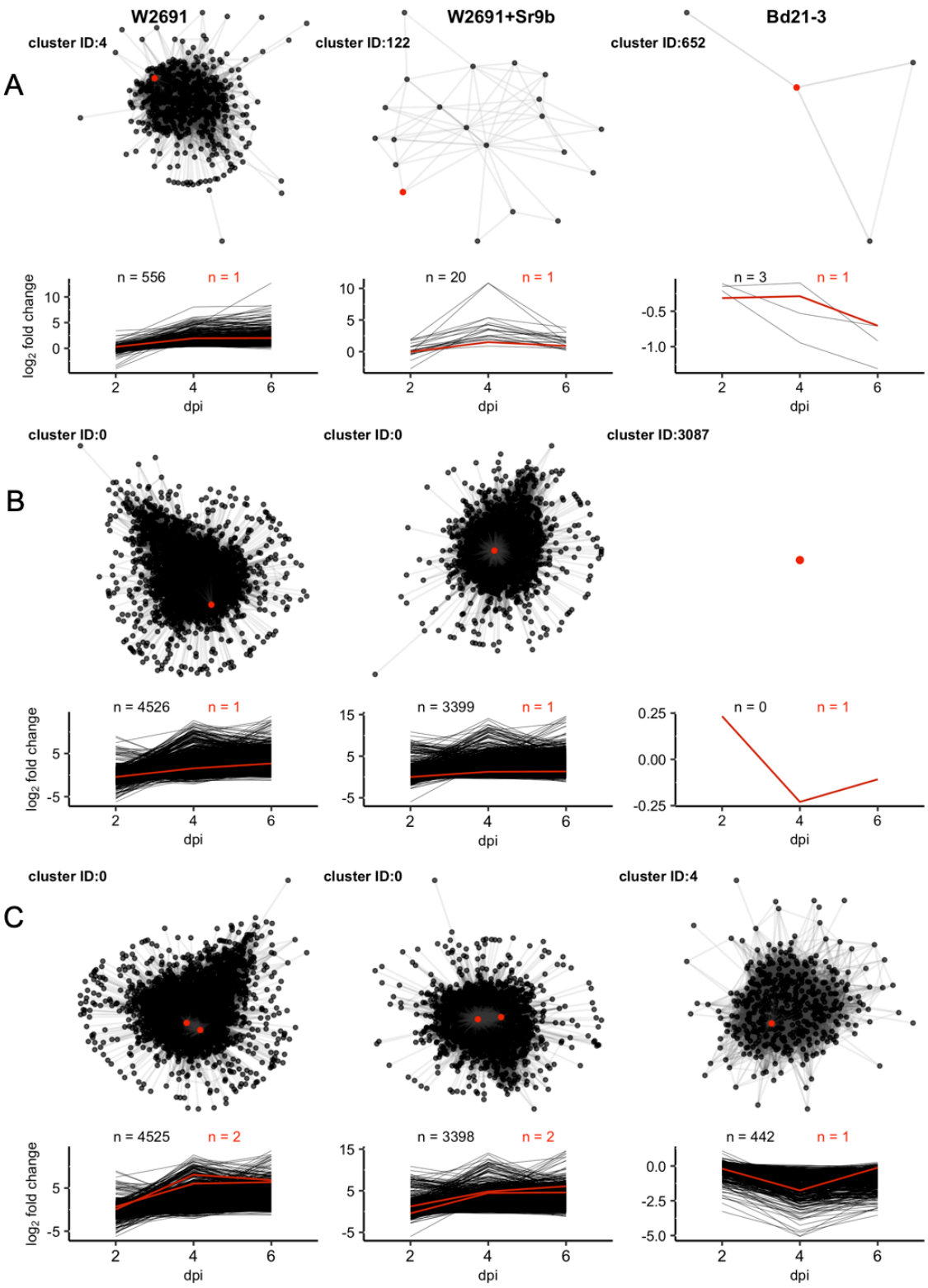



**Figure 2.6.** Network diagrams for clusters containing orthologs of **(A)** *DND1*, **(B)** *VAD1*, and (**C**) *DMR6* with corresponding plots showing log2 fold change of all nodes across 2,

4, and 6 dpi. Only connections with Z >= 3 are shown. Red lines, points, and counts represent *T. aestivum* and *B. distachyon* orthologs of *S* genes. Cluster identifiers (IDs) and gene names presented, left to right: *DND1:* 4 (TraesCS5D02G404600), 122 (TraesCS5D02G404600), 652 (BdiBd21-3.1G0110600); *VAD1*: 0 (TraesCS2D02G236800), 0 (TraesCS2D02G236800), 3085 (BdiBd21-3.1G0357000); *DMR6*: 0 (TraesCS4B02G346900), 0 (TraesCS4D02G341800), 4 (BdiBd21-3.1G1026800).

# Chapter 3: Towards a fully-phased whole genome assembly of a historic isolate of the oat crown rust fungus *Puccinia coronata* f. sp. *avenae*

**Contributions**

This chapter is in preparation for publication and has several collaborators who contributed work. I contributed:

- DNA and RNA extractions, preparation of spores for Hi-C preparation
- Initial Canu assembly and polishing as described below
- Assembly cleaning, telomere identification, and collapsed region analysis
- Assessment of isolate purity based on an allele frequency plot
- Alignment of bins to *Pgt* 21-0 A genome and visualization

---

Author contributions: Melania Figueroa, Peter Dodds, Jana Sperschneider and Shahryar F. Kianian conceived and designed the study, Eva C. Henningsen, Sheshanka Dugyala, Eric Nazareno, and Jakob Riddle and Roger Caspers conducted experiments, Eva C. Henningsen and Jana Sperschneider contributed to data analysis. Eva C. Henningsen, Jana Sperschneider, Melania Figueroa, Peter Dodds, Shahryar Kianian, and Brian J. Steffenson interpreted results. Eva C. Henningsen, Jana Sperschneider, Melania Figueroa, and Brian J. Steffenson wrote the manuscript; all authors contributed to manuscript editing and revisions.

## 3.1 Introduction

The biotrophic rust fungus *Puccinia coronata* f. sp. *avenae* (*Pca*) is the most damaging foliar pathogen of oat (*Avena sativa* L.) (Nazareno et al. 2018). *Pca* can destroy up to 50% of the crop during epidemics (USDA-ARS CDL 2014). *Pca* is commonly found in North and South America, Australia, and many other areas of the world where oats are grown (Nazareno et al. 2018). Like other pathogens *Pca* produces effectors, which help to overcome host defenses; however, as described by the gene-for-gene model and basic concepts in plant immunity, oat varieties have resistance (R) genes that encode immunoreceptors which recognize these effectors to stop the infection (Dodds and Rathjen 2010). Oat varieties which carry R genes are commonly deployed to control *Pca*, but due to rapid virulence evolution of the pathogen, epidemics still hinder grain production (Leonard and Martinelli 2005; Hammami et al. 2010). High quality genome references are essential to investigate the underlying mechanisms of virulence in pathogens and the identification of Avirulence (Avr) effectors (Figueroa et al. 2016, 2020). However, such resources have been difficult to generate for rust fungi given that these organisms are dikaryotic and have a complex genome structure. The genetic information of rust fungi is present in haplotypes physically separated in two nuclei for most of their life cycles; therefore, virulence phenotypes obey diploid genetics (Figueroa et al. 2016, 2020).

Until recently, genome references built for rust fungi did not fully capture the information from both nuclei and most assemblies available to researchers are haploid representations of the genome with abundant haplotype sequence swaps (Cantu et al. 2011; Cuomo et al. 2017). New technologies for generating long reads such as PacBio sequencing provided the first opportunities to attempt haplotype resolution in rust fungi (Miller et al. 2018;

Schwessinger et al. 2018). To date there are two *Pca* genome assemblies available, one for isolate 12SD80 and the other for isolate 12NC29 (Miller et al. 2018). These isolates were originally chosen for their distinct virulence profiles to support the identification of avirulence (Avr) effectors. However, it was recently documented that the population of *Pca* in the USA has changed over time (Miller et al. 2021), and therefore genome references from historic isolates may serve as excellent sources for Avr effector identification. While the genome assemblies provide valuable insights into the sequence divergence of haplotypes, assessments were limited by the persistence of collapsed haplotype regions in the references. State of the art *de novo* genome assembly techniques for rust fungi integrate PacBio reads and Hi-C data to reconstruct the entire sequence of chromosomes (Li et al. 2019), thus providing a unique opportunity to capture intra-organismal genetic diversity.

We selected one historic *Pca* isolate, named 203, to apply these techniques and as a foundation to generate a chromosome level genome assembly for the species that allows evaluation of haplotype divergence. This isolate is of interest in connection to former devastating epidemics of *Cochliobolus victoriae* on Victoria oats, a cultivar bred and deployed for its resistance to the prevalent *Pca* isolate 203 in the U.S. during the 1940's (Murphy and Meehan 1946). The resistance to 203 provided by the *Pc2* gene in Victoria is closely linked or is the same gene which contributes sensitivity to the *C. victoriae* toxin.

To generate a high-quality, fully haplotype-phased genome assembly for *Pca* isolate 203, high molecular weight DNA was extracted and sequenced. Hi-C data was collected to separate haplotypes and scaffold contigs, and RNA was collected at two timepoints (2 dpi and 5 dpi) for gene model prediction and annotation. Finally, the finished assembly was

compared with the previously published 12SD80 and 12NC29 assemblies. This assembly is expected to aid in future development of genomic resources for *Pca* and other rust species as well as the identification of avirulence effectors.

## 3.2 Materials and Methods

### 3.2.1 Plant and fungal materials and plant inoculations

The *Pca* isolate 203 (pathotype BBBGBCGLLB), known to be avirulent to the oat cultivar Victoria (Chang and Sadanaga 1964), was provided by the USDA-ARS Cereal Disease Laboratory, Saint Paul, MN, U.S.A. From this culture, a single pustule was isolated and increased to ensure isolate purity. Validation of the physiological race was conducted according to a standard nomenclature system using a set of oat differentials (Chong et al. 2000; Nazareno et al. 2018). Urediniospore stocks were kept at -80° C. As previously described by Miller et al. (2018), virulence phenotypes of *Pca* on oat differentials were converted to a 0-9 numerical scale, respectively, for heat map generation using R packages ComplexHeatmap and circlize (Gu et al. 2016, 2014). Seed from the oat differential set was also obtained from the Cereal Disease Laboratory. Spore increases were completed on the oat cultivar Marvelous as a universally susceptible host. Oat inoculations with *Pca* were carried out as described by Omidvar et al. (2018).

### 3.2.2 DNA or RNA isolation and sequencing

High molecular weight DNA was extracted from 700 mg of urediniospores increased from a single pustule of *Pca* 203 as previously described by Li et al. (2019) and sent to the University of Minnesota Genomics Center (UMGC, St. Paul, MN, U.S.A.) for library

construction using the PacBio SMRTbell 1.0 kit and sequencing using five PacBio Sequel System SMRT cells with v3 chemistry. DNA was also extracted from 20 mg of isolate 203 spores using the Omniprep$^{TM}$ DNA isolation kit from G-Biosciences for library preparation with the Illumina TruSeq Nano DNA protocol and Illumina NovaSeq sequencing in the S2 flow cell at UMGC to produce 150 bp paired-end reads. In preparation for Hi-C library preparation, 100 mg of spores were suspended in 1% formaldehyde and incubated at room temperature (RT) for 20 minutes with periodic vortexing. Glycine was added to this at 1g/100mL, and the suspension was incubated again at RT for 20 minutes with periodic vortexing. The suspension was centrifuged at 1000*g* for 1 minute and the supernatant was removed; the spores were then transferred to a liquid nitrogen-cooled mortar and ground before being stored at -80° C or on dry ice. Crosslinked spores for Hi-C library preparation with the Phase Genomics Proximo Fungal 4.0 protocol were sent to Phase Genomics (Seattle, WA, U.S.A), and libraries were sequenced to 100 million 150 bp paired-end reads at Genewiz (South Plainfield, NJ, U.S.A.). Finally, RNA was extracted from infected oat at 2 and 5 days post-inoculation (dpi) using the Qiagen RNeasy Plant Mini Kit following the manufacturer's instructions and sent to UMGC for library preparation with the Illumina TruSeq Stranded mRNA protocol and sequencing using Illumina NextSeq in mid-output mode, producing 75 bp paired-end reads.

### 3.2.3 Genome assembly and polishing

Isolate purity was confirmed by calling SNPs against the existing 12SD80 genome assembly and calculating allele frequencies (Appendix G) (Miller et al. 2018). For this, Illumina short reads were trimmed with trimmomatic version 0.33 (Bolger et al. 2014) and

aligned with bwa version 0.7.17 (Li and Durbin 2009). The alignments were prepared with samtools version 1.9 (Li et al. 2009), and variants were called with Freebayes version 1.1.0 (Garrison and Marth 2012). The allele balance plot was generated using a custom R script (https://github.com/henni164/Pca203_assembly/figure_s1/203_frequencies.R).

An initial genome assembly was constructed from 2,260,466 subreads using Canu version 2.1 (Koren et al. 2017). Half of the subreads (2/4 subread bam files) were mapped back to the assembly using PacBio software pbmm2 version 1.4.0 (https://github.com/PacificBiosciences/pbmm2/releases/tag/v1.4.0) and a new consensus generated with PacBio software GenomicConsensus version 2.3.3 (https://github.com/PacificBiosciences/GenomicConsensus/releases/tag/2.3.3). This process was repeated using the new consensus as the reference. Next, the updated consensus was polished twice in Pilon version 1.22 using 46,950,034 Illumina short reads trimmed with trimmomatic version 0.33 (Bolger et al. 2014; Walker et al. 2014). Assembly statistics were calculated using Quast (Gurevich et al. 2013).

### 3.2.4 Identification of mitochondrial contigs and removal of assembly contaminants or artifacts

All contigs were first blasted to the mitochondrial genome database from NCBI with ncbiblast+ version 2.8.1, and mitochondrial contigs were removed from the main assembly (Camacho et al. 2009). The remaining contigs were blasted to the NCBI nucleotide library using ncbiblast+ v.2.8.1 and other likely contaminants were removed (Camacho et al. 2009). Finally, contig coverage was also examined; contigs with coverage <2x were removed as they likely represent contaminants. Telomeres and collapsed regions were

identified with custom scripts (https://github.com/JanaSperschneider/GenomeAssemblyTools, https://github.com/JanaSperschneider/FindTelomeres).

### 3.2.5 Curation, contig binning, and haplotype-phasing of genome assembly

To generate a clean assembly, a table of BUSCO gene hits (BUSCO 3.1.0 -l basidiomycota_odb9 -m geno -sp coprinus) (Simão et al. 2015) was produced as well as a table of gene hits from the *Puccinia coronata* f. sp. *avenae* 12SD80 transcript set (Miller et al. 2018) with biokanga blitz (4.4.2 --sensitivity=2 --mismatchscore=1) (https://github.com/csiro-crop-informatics/biokanga). Only genes that have exactly two hits were retained as phasing markers. All-versus-all contig alignments were computed with minimap2 (-k19 -w19 -m200 -DP -r1000) to find syntenous contigs (Li 2018). We used the duplicated gene information to put contigs that share genes into scaffold bins. For each possible pair of contigs, we recorded the number of their shared genes (number of shared BUSCO genes + number of shared *Pca* genes based on previous annotations (Miller et al. 2018)). The total number of shared genes was normalized to shared gene density per Mb. We then constructed a graph where each contig is a node and a pair of contigs is connected by a weighted edge comprising the shared gene density. Two contigs were connected by an edge if their shared gene density per Mb is greater than 30, if they share more than two genes and if one of the contigs has more than 20% of its bases aligned with the other. A graph network approach was then used to identify scaffold bins that contain homologous pairs of sequences from each haplotype (Python's NetworkX community.best_partition) (Blondel et al. 2008). Scaffold bins that contain contigs with a

combined size > 1Mb were kept and for each scaffold bin *x*, the contigs within were separated into the two haplotype sets. To assist in the identification of adjacent bins, we relied on synteny with the related species *Puccinia graminis* f. sp. *tritici* (*Pgt*) (Li et al. 2019). Bins were aligned to the *Pgt* isolate 21-0 genome assembly of haplotype A with promer (mummer version 3.23) to identify likely adjacent bins (Kurtz et al. 2004; Li et al. 2019). The alignments were visualized using a custom R script (https://github.com/henni164/Pca203_assembly/dotplots.R).

### 3.2.6 Data Availability

The completed assembly and annotation will be available on NCBI. Scripts related to identification of contaminants, collapsed regions, and telomeres are available at https://github.com/JanaSperschneider/GenomeAssemblyTools and https://github.com/JanaSperschneider/FindTelomeres. Scripts and data used to construct figures are available at https://github.com/henni164/Pca203_assembly.

## 3.3 Results

### 3.3.1 Virulence profile and pathotyping of *Pca* isolate 203

Given that the virulence profile of isolate *Pca* 203 has never been examined in the expanded oat differential set that is commonly used in North America (Nazareno et al. 2018), a full report of infection scores for this isolate was generated (Figure 3.1). In contrast to the *Pca* isolates (12NC29 and 12SD80) for which genome references have been constructed, *Pca* 203 is more avirulent than the other two isolates as illustrated by the heatmap. These results indicate that *Pca* 203 contains sequences for Avr effectors that are

absent in the isolates 12NC29 and 12SD80 and can therefore assist in effector identification and studies of virulence evolution for resistance genes represented in the oat differential set.

### 3.3.2 Genome reference description

According to previous reports, the genome size of *Pca* ranges from 99.16 to 105.25 Mb (Miller et al. 2018). In total, 51 Gb of PacBio data (~100X coverage of the *Pca* genome) and 19 Gb of Illumina short reads (~58X coverage of the *Pca* genome) was acquired to generate a *de novo* genome draft of the isolate *Pca* 203. Using these data, a raw genome assembly was constructed from 2,260,466 PacBio subreads which resulted in a total length of 206,897,753 bp and spanned 678 contigs (Table 3.1). A total of 542 contigs were larger than 50 Kb and added to 202,442,886 bp. Considering the estimated size of the *Pca* genome (~100 Mb), these results suggest that most of the information from both nuclei was captured by this approach. Genome polishing with Illumina short reads was conducted two times as previously optimized (Li et al. 2019) and corrected 6,371 indels in the first round and 229 indels in the second round. Once the initial genome assembly was polished, a search of mitochondrial contigs was performed. Three mitochondrial contigs were identified and removed from the main assembly (tig494, tig848, and tig849). In addition to this, contigs were inspected for coverage by using BBmap (https://sourceforge.net/projects/bbmap/). The average genome coverage across contigs was 117X, therefore a <2X coverage threshold was selected to yield a clean assembly. This resulted in seven contigs with <2X coverage being removed (tig397, tig523, tig719, tig718, tig726, tig739, and tig777). Finally, a BLAST search of the NCBI database revealed seven

contaminant contigs (tig492, tig493, tig505, tig506, tig507, tig852, and tig853) and one PacBio construct (tig496). Later investigation of telomeric sequences identified one mis-assembly contig containing repeat reads, which was also removed (tig504). In total, 20 contigs were removed from the final assembly (Appendix G). This resulted in a cleaned assembly with 658 covering 206,392,043 bp with a contig N50 size of 798,317 bp (Table 3.1).

Collapsed regions were determined by plotting a histogram of the average coverage of 1000 bp bins (Figure 3.2). The first large peak represents the coverage for most of the haploid or non-collapsed regions, which is 117X coverage. A second small peak near 224X coverage, or double the haploid coverage, would likely represent collapsed regions. To capture the likely full range of collapsed regions, any area with coverage greater than (haploid peak + (diploid peak – haploid peak/2)) were considered collapsed. In this case that value was 170X coverage, and from this value there were 2.259 Mb of collapsed regions in the assembly. These regions were distributed across 126 contigs, with the smallest frequency per contig being one and the highest being 31 on tig549.

### 3.3.3 Assessing genome completeness

Telomere sequence analysis identified 66 telomeres on the end of contigs, with 3 contigs containing both telomeres (tig00000236, tig000000330, and tig000000410 at 4.7, 4.3, and 3.41 Mbp long, respectively). Initial contig binning with 12SD80 RNAseq data (Miller et al. 2018) yielded 45 bins with 413 contigs, which represented 183.062641 Mb of the genome (Appendix G). Bins were aligned to the genome reference of Pgt21-0 haplotype A genome (Li et al. 2019) to determine possible relationships between bins; each

chromosome had between 1-4 bins in clear alignment (Appendix G). Contigs with telomeres were checked to ensure that each collection of bins had an appropriate number of telomeres; sets of bins representing putative chromosomes had between zero and two chromosomes per haplotype, indicating that bin alignment to Pgt21-0 did not result in inappropriate bin grouping (i.e. 2> telomeres in one chromosome group). Adding together the total length of all contigs that aligned to each *Pgt* chromosome allowed some preliminary knowledge of the lengths of the *Pca* chromosomes, which is summarized in Appendix G.

### 3.3.4 Next steps

A fully-phased, chromosome level assembly of *Pca* 203 is not yet complete. Currently, manual curation is underway to repair haplotype swaps. Contact maps generated with HiC-Pro (Servant et al. 2015) and HiCExplorer (Wolff et al. 2020) were used to identify likely regions of haplotype swaps. Illumina reads were mapped to the cleaned genome assembly using bwa (Li and Durbin 2009) and visualized using the online IGV tool (Thorvaldsdóttir et al. 2013). Candidate breakpoints will be determined using breaks in read coverage or SNPs, and then contigs will be broken and re-arranged. Contact maps will be generated again to determine if the changes resulted in accurate haplotype assignments. Scaffolding to generate chromosome pseudomolecules will be performed with SALSA2 (Ghurye et al. 2019). Following this, the phased, chromosome-level assembly will be annotated. In preparation for gene model prediction, the assembly will be repeat-masked with repeatmasker version 4.1.1 using a custom library generated with repeatmodeler version 2.0.1 (Smit et al. 2013; Flynn et al. 2020). The RNAseq reads will be processed by

trimming with trimmomatic version 0.33 (Bolger et al. 2014). Reads will be aligned to the assembly with hisat2 version 2.1.0 and transcripts will be assembled with Trinity version 2.5.1 (Grabherr et al. 2011; Kim et al. 2019). The assembly will be annotated with trinity transcripts using by Funnanotate version 1.8.1 (Palmer and Stajich 2020).

## 3.4 Discussion

So far, only a few genome references for rust fungi have been assembled to fully represent haplotype content. Here, a high-quality *de novo* draft genome assembly and annotation of a historic isolate of the oat crown rust fungus are in progress and will enable future pathogenicity studies of this important pathogen that hinders oat production in many parts of the world. The assessments of genome completeness (genome size, collapsed regions, telomere identification) indicate that the assembly process collapsed a negligible amount of the genome and that a full or near-full genome representation is present; 206.4 Mbp covered by 658 contigs have been assembled, and only about 2 Mbp of this is considered collapsed. Even in a state of partial completion, the assembly was less fragmented and more complete than previous assemblies for *Pca* isolates 12SD80 (150.47 Mbp, 603 primary contigs, 1,033 haplotigs) and 12NC29 (166.28 Mbp, 777 primary contigs, 950 haplotigs) (Miller et al. 2018). The alignment of 45 haplotype bins to the *Pgt* 21-0 A genome indicates that there are likely 18 chromosomes, and that the bins represent close to the full chromosome length. This is in range with previous predictions indicating that *Pca* has 16-20 chromosomes (Boehm 1992). A preliminary review of the total length of the bins aligning to each *Pgt* chromosome indicate the shortest chromosome is around 2.9 Mb and the longest is around 7.4 Mb; results similar to those reported in *Pgt* (Li et al. 2019).

The next steps to complete the assembly are fixing phase swaps between haplotypes, scaffolding contigs and assigning chromosomes, and annotating the completed genome. Completion of this reference has the potential to fuel novel research and to solve old questions left unanswered. With more high-quality references becoming available for diverse rust fungi, it will soon be possible to conduct comparative genomics studies within species and even between rust genera. Questions such as whether effectors are conserved between rusts, the genetic relationship between rusts, and whether rusts exchange genetic material can be explored at a much finer resolution than ever before. Old questions may be answered too. For example, it has still not been resolved whether *Pc2* (*Pca* resistance gene 2) and *Vb* (victorin receptor) are the same gene or not. The toxin produced by *C. victoriae* was identified and a putative receptor in oat was characterized, but no further information is known (Wolpert et al. 1985; Wolpert and Macko 1989). Breeding efforts demonstrated that *Pc2* and Vb are most likely the same gene, as no individuals segregating for Victoria-type resistance and victorin sensitivity were recovered (Welsh et al. 1954; Mayama et al. 1995). Later work in Arabidopsis demonstrated that the *A. thaliana* victorin receptor LOV1 is also an NBS-LRR (nucleotide binding site leucine-rich repeat), indicating that *Pc2* could also have such a pleiotropic effect and cause susceptibility to *C. victoriae*; however, this has not actually been confirmed (Lorang et al. 2007). The LOV1 homeolog in *H. vulgare*, *HvLov1*, is also being examined and research is underway to determine the linkage to *Mla3*, as well as the sensitivity to victorin conferred by the locus (Brabham et al. 2019). The creation of a high-quality assembly of isolate 203, which should contain *AvrPc2* and therefore is recognized by *Pc2*, may provide a clue to solving this decades-long mystery.

**Table 3.1.** Statistics for the raw and cleaned *Puccinia coronata* f. sp. *avenae* isolate 203 assemblies.

| **Assembly statistic** | **Raw** | **Cleaned** |
|---|---|---|
| # of contigs | 678 | 658 |
| # of contigs >= 0 bp | 678 | 658 |
| # of contigs >= 1000 bp | 678 | 658 |
| # of contigs >= 5000 bp | 670 | 655 |
| # of contigs >= 10000 bp | 668 | 655 |
| # of contigs >= 25000 bp | 664 | 654 |
| # of contigs >= 50000 bp | 542 | 539 |
| Largest contig (bp) | 5,222,385 | 5,222,753 |
| Total length >= 0 bp | 206,897,753 | 206,392,043 |
| Total length >= 1000 bp | 206,897,753 | 206,392,043 |
| Total length >= 5000 bp | 206,876,127 | 206,384,417 |
| Total length >= 10000 bp | 206,861,036 | 206,384,417 |
| Total length >= 25000 bp | 206,791,090 | 206,359,517 |
| Total length >= 50000 bp | 202,442,886 | 202,224,731 |
| N50 | 785,932 | 798,317 |
| N75 | 266,072 | 266,628 |
| L50 | 54 | 53 |
| L75 | 181 | 179 |
| GC% | 44.63 | 44.65 |
| Number of N's per 100 kbp | 0.00 | 0.00 |

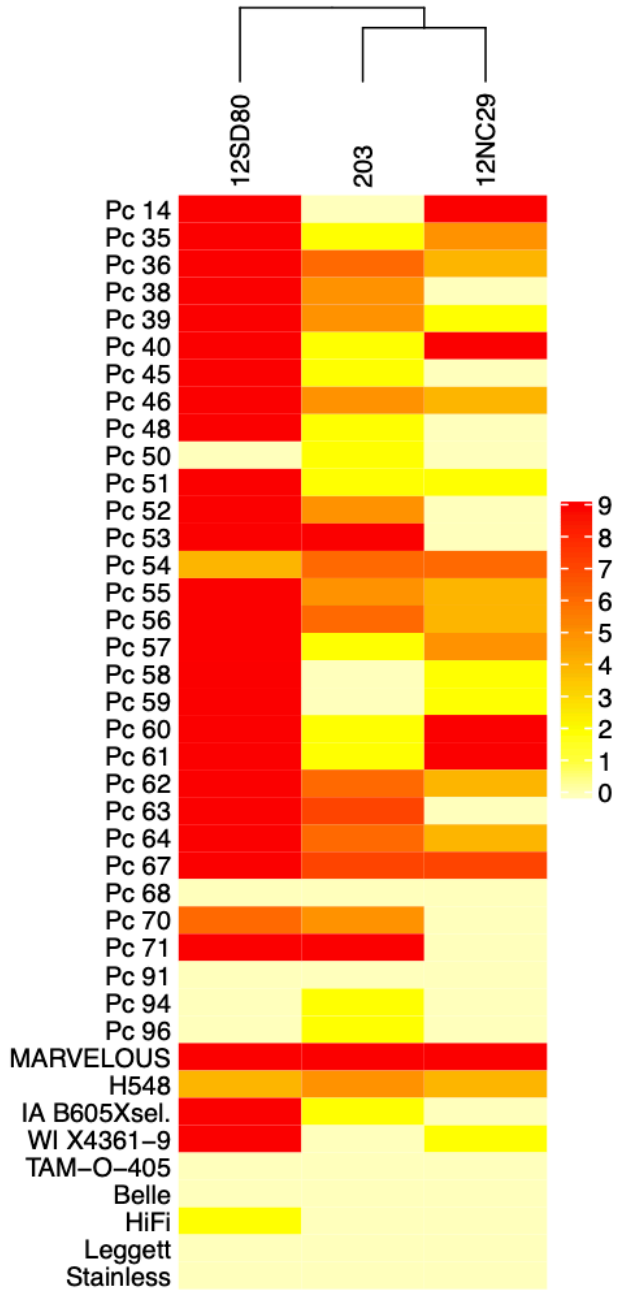



**Figure 3.1**. Heatmap of linearized rust scores for *Puccinia coronata* f. sp. *avenae* isolates 203, 12NC29, and 12SD80 on the North American oat differential set. Description of 12NC29 and 12SD80 is shown by Miller et al. 2018. Infection scores were converted to a numeric scale (0 = resistance to 9 = susceptibility).

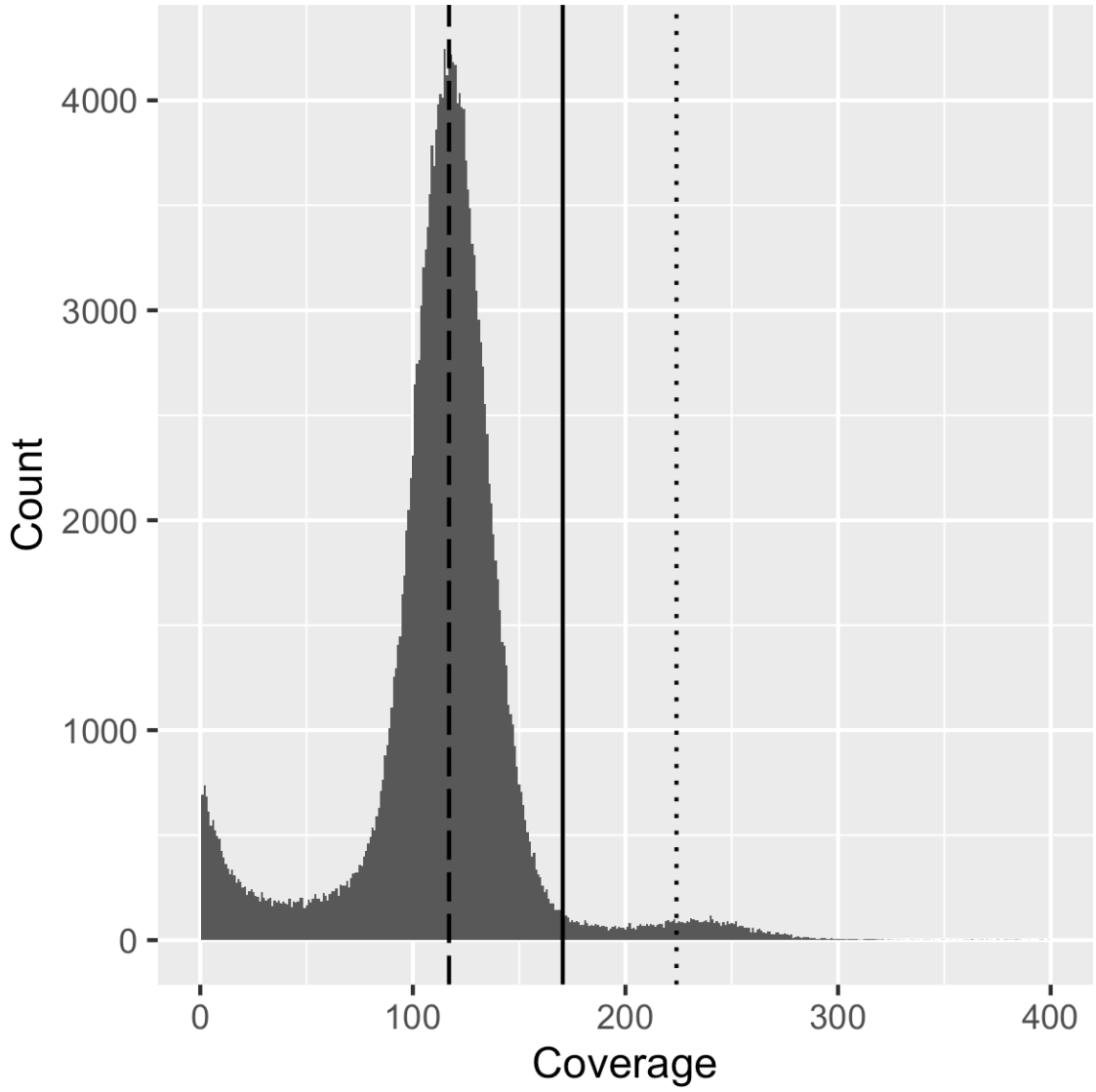



**Figure 3.2**. Histogram of the average coverage in 1000 bp bins across the cleaned *Pca* 203 genome assembly. Dashed line (x = 117) represents the coverage mode, which is likely the haploid coverage for the assembly. The dotted line represents double the coverage mode, or double haploid coverage (x = 224). The solid line (x = 170) is the cutoff used for determining whether a region was considered collapsed.

# Bibliography


2014 Wheat Stem Rust Observations in the U.S. 2014. St. Paul, Minn.: USDA-ARS Cereal Disease Laboratory. Available at: https://www.ars.usda.gov/ARSUserFiles/50620500/Cerealrustbulletins/2014wsr.pdf

Acevedo-Garcia, J., Spencer, D., Thieron, H., Reinstädler, A., Hammond-Kosack, K., Phillips, A. L., et al. 2017. *mlo*-based powdery mildew resistance in hexaploid bread wheat generated by a non-transgenic TILLING approach. Plant Biotechnol. J. 15:367–378.

Admassu-Yimer, B., Bonman, J. M., and Esvelt Klos, K. 2018. Mapping of crown rust resistance gene *Pc53* in oat (*Avena sativa*). PLOS ONE. 13:e0209105.

Ai, G., Yang, K., Ye, W., Tian, Y., Du, Y., Zhu, H., et al. 2020. Prediction and Characterization of RXLR Effectors in *Pythium* Species. Mol. Plant-Microbe Interactions. 33:1046–1058.

Aime, M. C., McTaggart, A. R., Mondo, S. J., and Duplessis, S. 2017. Chapter Seven – Phylogenetics and Phylogenomics of Rust Fungi. In *Advances in Genetics*, eds. Jeffrey P. Townsend and Zheng Wang. Academic Press, p. 267–307. Available at: https://www.sciencedirect.com/science/article/pii/S0065266017300391.

Alaux, M., Rogers, J., Letellier, T., Flores, R., Alfama, F., Pommier, C., et al. 2018. Linking the International Wheat Genome Sequencing Consortium bread wheat reference genome sequence to wheat genetic and phenomic data. Genome Biol. 19:111.

Alexa, A., and Rahnenfuhrer, J. 2020. *topGO: Enrichment Analysis for Gene Ontology*.

Anders, S., Pyl, P. T., and Huber, W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinforma. Oxf. Engl. 31:166–169.

Ayliffe, M., Singh, D., Park, R., Moscou, M., and Pryor, T. 2013. Infection of *Brachypodium distachyon* with selected grass rust pathogens. Mol. Plant-Microbe Interact. MPMI. 26:946–957.

Berghaus, R., and Reisener, H. J. 1984. Changes in photosynthesis of wheat plants infected with wheat stem rust (*Puccinia graminis* f. sp. *tritici*). Phytopathology. 112:165–172.

Bettgenhaeuser, J., Gardiner, M., Spanner, R., Green, P., Hernández-Pinzón, I., Hubbard, A., et al. 2018. The genetic architecture of colonization resistance in *Brachypodium distachyon* to non-adapted stripe rust (*Puccinia striiformis*) isolates. PLoS Genet. 14.

Bettgenhaeuser, J., Gilbert, B., Ayliffe, M., and Moscou, M. J. 2014. Nonhost resistance to rust pathogens-a continuation of continua. Front. Plant Sci. 5:664.

Bhattacharya, S. 2017. Deadly new wheat disease threatens Europe's crops. Nature. 542(7640):145–146.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. 2008. Fast unfolding of communities in large networks. J. Stat. Mech. Theory Exp. 2008:P10008.

Boehm, E. W. A. 1992. Determination of the karyotype in selected species of *Eocronartium*, *Puccinia*, and *Melampsora*.

Bolger, A. M., Lohse, M., and Usadel, B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics.

Bouchez, O., Huard, C., Lorrain, S., Roby, D., and Balagué, C. 2007. Ethylene is one of the key elements for cell death and defense response control in the *Arabidopsis* lesion mimic mutant *vad1*. Plant Physiol. 145:465.

Bozkurt, T. O., and Kamoun, S. 2020. The plant–pathogen haustorial interface at a glance. J. Cell Sci. 133:jcs237958.

Bozkurt, T. O., McGrann, G. R., MacCormack, R., Boyd, L. A., and Akkaya, M. S. 2010. Cellular and transcriptional responses of wheat during compatible and incompatible race-specific interactions with *Puccinia striiformis* f. sp. *tritici*. Mol. Plant Pathol. 11:625–640.

Brabham, H.J., Hernández-Pinzón, I., Lorang, J.M., Wolpert, T.J., Hayes, P.M., Sato, K., et al. 2019, July 14-18. Multiple pathogen recognition by *Mla3* in barley. International Society for Molecular Plant-Microbe Interactions XVIII Congress. Glasgow, Scotland, United Kingdom. https://ismpmi.confex.com/ismpmi/2019/meetingapp.cgi/Paper/4242

Bragg, J. N., Wu, J., Gordon, S. P., Guttman, M. E., Thilmony, R., Lazo, G. R., et al. 2012. Generation and characterization of the Western Regional Research Center *Brachypodium* T-DNA insertional mutant collection. PLOS ONE. 7:e41916.

Briatte, F. 2020. *ggnetwork: Geometries to Plot Networks with "ggplot2."* Available at: https://CRAN.R-project.org/package=ggnetwork.

Büschges, R., Hollricher, K., Panstruga, R., Simons, G., Wolter, M., Frijters, A., et al. 1997. The Barley *Mlo* Gene: A novel control element of plant pathogen resistance. Cell. 88:695–705.

Bushnell, W. R. 1984. 15 - Structural and Physiological Alterations in Susceptible Host Tissue. In *The Cereal Rusts*, eds. William R. Bushnell and Alan P. Roelfs. Academic Press, p. 477–507. Available at: http://www.sciencedirect.com/science/article/pii/B9780121484019500212.

Butts, C. T. 2015. *network: Classes for Relational Data*. Available at: https://CRAN.R-project.org/package=network.

Butts, C. T. 2019. *sna: Tools for Social Network Analysis*. Available at: https://CRAN.R-project.org/package=sna.

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

Cantu, D., Govindarajulu, M., Kozik, A., Wang, M., Chen, X., Kojima, K. K., et al. 2011. Next generation sequencing provides rapid access to the genome of *Puccinia striiformis* f. sp. *tritici*, the causal agent of wheat stripe rust. PLOS ONE. 6:e24230.

Cantu, D., Segovia, V., MacLean, D., Bayles, R., Chen, X., Kamoun, S., et al. 2013. Genome analyses of the wheat yellow (stripe) rust pathogen *Puccinia striiformis* f. sp. *tritici* reveal polymorphic and haustorial expressed secreted proteins as candidate effectors. BMC Genomics. 14:270.

Chandra, S., Singh, D., Pathak, J., Kumari, S., Kumar, M., Poddar, R., et al. 2016. *De novo* assembled wheat transcriptomes delineate differentially expressed host genes in response to leaf rust. PLoS ONE. 11.

Chang, T. D., and Sadanaga, K. 1964. Crosses of six monosomics in *Avena sativa* L. with varieties, species, and chlorophyll mutants. Crop Sci. 4:cropsci1964.0011183X000400060012x.

Chen, B., Jiang, J., and Zhou, X. 2007. A TOM1 homologue is required for multiplication of Tobacco Mosaic Virus in *Nicotiana benthamiana*. J. Zhejiang Univ. Sci. B. 8:256–259.
Chen, J., Upadhyaya, N. M., Ortiz, D., Sperschneider, J., Li, F., Bouton, C., et al. 2017. Loss of *AvrSr50* by somatic exchange in stem rust leads to virulence for *Sr50* resistance in wheat. Science. 358:1607.
Chen, J., Wu, J., Zhang, P., Dong, C., Upadhyaya, N. M., Zhou, Q., et al. 2019. *De Novo* genome assembly and comparative genomics of the barley leaf rust pathogen *Puccinia hordei* identifies candidates for three avirulence genes. G3 GenesGenomesGenetics. 9:3263.
Chen, L.-Q., Hou, B.-H., Lalonde, S., Takanaga, H., Hartung, M. L., Qu, X.-Q., et al. 2010. Sugar transporters for intercellular exchange and nutrition of pathogens. Nature. 468:527–532.
Chong, J., Leonard, K. J., and Salmeron, J. J. 2000. A North American system of nomenclature for *Puccinia coronata* f. sp. *avenae*. Plant Dis. 84:580–585.
Clough, S. J., Fengler, K. A., Yu, I. C., Lippok, B., Smith, R. K., Jr, and Bent, A. F. 2000. The *Arabidopsis dnd1* "defense, no death" gene encodes a mutated cyclic nucleotide-gated ion channel. Proc. Natl. Acad. Sci. U. S. A. 97:9323–9328.
Couto, D., and Zipfel, C. 2016. Regulation of pattern recognition receptor signalling in plants. Nat. Rev. Immunol. 16:537–552.
Cuomo, C. A., Bakkeren, G., Khalil, H. B., Panwar, V., Joly, D., Linning, R., et al. 2017. Comparative analysis highlights variable genome content of wheat rusts and divergence of the mating loci. G3 GenesGenomesGenetics. 7:361–376.
Ding, P., and Ding, Y. 2020. Stories of salicylic acid: A plant defense hormone. Trends Plant Sci. 25:549–565.
Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. 2012. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 29:15–21.
Dobon, A., Bunting, D. C., Cabrera-Quio, L. E., Uauy, C., and Saunders, D. G. 2016. The host-pathogen interaction between wheat and yellow rust induces temporally coordinated waves of gene expression. BMC Genomics. 17:380.
Dodds, P. N., and Rathjen, J. P. 2010. Plant immunity: towards an integrated view of plant-pathogen interactions. Nat. Rev. Genet. 11:539–548.
Duplessis, S., Cuomo, C. A., Lin, Y. C., Aerts, A., Tisserant, E., Veneault-Fourrey, C., et al. 2011. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. Proc. Natl. Acad. Sci. 108:9166–9171.
Ellis, J. G., Lagudah, E. S., Spielmeyer, W., and Dodds, P. N. 2014. The past, present and future of breeding rust resistant wheat. Front. Plant Sci. 5:641.
Emms, D. M., and Kelly, S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 20:238.
Engelhardt, S., Stam, R., and Hückelhoven, R. 2018. Good riddance? Breaking disease susceptibility in the era of new breeding technologies. Agronomy. 8.
Eulgem, T., and Somssich, I. E. 2007. Networks of WRKY transcription factors in defense signaling. Spec. Issue Biot. Interact. 10:366–371.
Figueroa, M., Alderman, S., Garvin, D. F., and Pfender, W. F. 2013. Infection of *Brachypodium distachyon* by formae speciales of *Puccinia graminis*: Early infection events and host-pathogen incompatibility. PLOS ONE. 8:e56857.

Figueroa, M., Castell-Miller, C. V., Li, F., Hulbert, S. H., and Bradeen, J. M. 2015. Pushing the boundaries of resistance: insights from *Brachypodium*-rust interactions. Front. Plant Sci. 6:558.
Figueroa, M., Dodds, P. N., and Henningsen, E. C. 2020. Evolution of virulence in rust fungi — multiple solutions to one problem. Curr. Opin. Plant Biol. 56:20–27.
Figueroa, M., Upadhyaya, N. M., Sperschneider, J., Park, R. F., Szabo, L. J., Steffenson, B., et al. 2016. Changing the game: using integrative genomics to probe virulence mechanisms of the stem rust pathogen *Puccinia graminis* f. sp. *tritici*. Front. Plant Sci. 7:205.
Flor, H. H. 1971. Current status of the gene-for-gene concept. Annu. Rev. Phytopathol. 9:275–296.
Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., et al. 2020. RepeatModeler2 for automated genomic discovery of transposable element families. Proc. Natl. Acad. Sci. 117:9451.
Gaj, T., Gersbach, C. A., and Barbas, C. F., 3rd. 2013. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. 31:397–405.
Garnica, D. P., Nemri, A., Upadhyaya, N. M., Rathjen, J. P., and Dodds, P. N. 2014. The ins and outs of rust haustoria. PLoS Pathog. 10.
Garrison, E., and Marth, G. 2012. Haplotype-based variant detection from short-read sequencing. ArXiv Prepr.
Ge, X., Deng, W., Lee, Z. Z., Lopez-Ruiz, F. J., Schweizer, P., and Ellwood, S. R. 2016. Tempered *mlo* broad-spectrum resistance to barley powdery mildew in an Ethiopian landrace. Sci. Rep. 6:29558.
Genger, R. K., Jurkowski, G. I., McDowell, J. M., Lu, H., Jung, H. W., Greenberg, J. T., et al. 2008. Signaling pathways that regulate the enhanced disease resistance of *Arabidopsis* "defense, no death" mutants. Mol. Plant-Microbe Interact. MPMI. 21:1285–1296.
Ghurye, J., Rhie, A., Walenz, B. P., Schmitt, A., Selvaraj, S., Pop, M., et al. 2019. Integrating Hi-C links with assembly graphs for chromosome-scale assembly. PLOS Comput. Biol. 15:e1007273.
Gilbert, B., Bettgenhaeuser, J., Upadhyaya, N., Soliveres, M., Singh, D., Park, R. F., et al. 2018. Components of *Brachypodium distachyon* resistance to nonadapted wheat stripe rust pathogens are simply inherited. PLoS Genet. 14.
Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29:644–652.
Gu, Z., Eils, R., and Schlesner, M. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 32:2847–2849.
Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. 2014. circlize implements and enhances circular visualization in R. Bioinformatics. 30:2811–2812.
Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 29:1072–1075.
Hammami, I., Allagui, M. B., Chakroun, M., and E-Gazzeh, M. 2010. Natural population of oat crown rust in Tunisia. Phytopathol. Mediterr. 49:35–41.
Harder, D. E., and Chong, J. 1984. Structure and physiology of haustoria. In *The Cereal Rusts*, ed. William Bushnell. Academic Press, p. 416–460.

Helliwell, C., and Waterhouse, P. 2003. Constructs and methods for high-throughput gene silencing in plants. RNA Interf. 30:289–295.
Henningsen, E., Sallam, A. H., Matny, O., Szinyei, T., Figueroa, M., and Steffenson, B. J. 2020. *Rpg7*: A new gene for stem rust resistance from *Hordeum vulgare* ssp. *spontaneum*. Phytopathology. 111:548–558.
Howe, K. L., Contreras-Moreira, B., De Silva, N., Maslen, G., Akanni, W., Allen, J., et al. 2019. Ensembl Genomes 2020—enabling non-vertebrate genomic research. Nucleic Acids Res. 48:D689–D695.
Huang, T.-Y., Desclos-Theveniau, M., Chien, C.-T., and Zimmerli, L. 2013. *Arabidopsis thaliana* transgenics overexpressing *IBR3* show enhanced susceptibility to the bacterium *Pseudomonas syringae*. Plant Biol. 15:832–840.
Huttenhower, C., Hibbs, M., Myers, C., and Troyanskaya, O. G. 2006. A scalable method for integration and functional analysis of multiple microarray datasets. Bioinformatics. 22:2890–2897.
Jia, H., Zhang, Y., Orbović, V., Xu, J., White, F. F., Jones, J. B., et al. 2017. Genome editing of the disease susceptibility gene *CsLOB1* in citrus confers resistance to citrus canker. Plant Biotechnol. J. 15:817–823.
Jørgensen, I. H. 1992. Discovery, characterization and exploitation of *Mlo* powdery mildew resistance in barley. Euphytica. 63:141–152.
Jurkowski, G. I., Smith, R. K., Yu, I., Ham, J. H., Sharma, S. B., Klessig, D. F., et al. 2004. *Arabidopsis DND2*, a second cyclic nucleotide-gated ion channel gene for which mutation causes the "defense, no death" phenotype. Mol. Plant-Microbe Interactions. 17:511–520.
Kellogg, E. A. 2001. Evolutionary history of the grasses. Plant Physiol. 125:1198–1205.
Khafif, M., Balagué, C., Huard-Chauveau, C., and Roby, D. 2017. An essential role for the VASt domain of the *Arabidopsis VAD1* protein in the regulation of defense and cell death in response to pathogens. PLOS ONE. 12:e0179782.
Kim, D., Alptekin, B., and Budak, H. 2018. CRISPR/Cas9 genome editing in wheat. Funct. Integr. Genomics. 18:31–41.
Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. 37:907–915.
Knight, K. S., Kurylo, J. S., Endress, A. G., Stewart, J. R., and Reich, P. B. 2007. Ecology and ecosystem impacts of common buckthorn (*Rhamnus cathartica*): a review. Biol. Invasions. 9:925–937.
Kolmer, J. A., Ordoñez, M. E., and Groth, J. V. 2018. The Rust Fungi. eLS.
Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. Genome Res. 27:722–736.
Kretschmer, M., Damoo, D., Djamei, A., and Kronstad, J. 2020. Chloroplasts and plant immunity: Where are the fungal effectors? Pathogens. 9.
Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., et al. 2004. Versatile and open software for comparing large genomes. Genome Biol. 5:R12.
Kurylo, J., and Endress, A. G. 2012. *Rhamnus cathartica*: Notes on its early history in North America. Northeast. Nat. 19:601–610.

Laflamme, B., Dillon, M. M., Martel, A., Almeida, R. N. D., Desveaux, D., and Guttman, D. S. 2020. The pan-genome effector-triggered immunity landscape of a host-pathogen interaction. Science. 367:763.
Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., et al. 2011. The *Arabidopsis* Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res. 40:D1202–D1210.
Lapin, D., and Van den Ackerveken, G. 2013. Susceptibility to plant disease: more than a failure of host immunity. Trends Plant Sci. 18:546–554.
Lee, W.-S., Rudd, J. J., and Kanyuka, K. 2015. Virus induced gene silencing (VIGS) for functional analysis of wheat genes involved in *Zymoseptoria tritici* susceptibility and resistance. Fungal Genet. Biol. 79:84–88.
Lemoine, F., Correia, D., Lefort, V., Doppelt-Azeroual, O., Mareuil, F., Cohen-Boulakia, S., et al. 2019. NGPhylogeny.fr: new generation phylogenetic services for non-specialists. Nucleic Acids Res. 47:W260–W265.
Leonard, K. J., and Martinelli, J. A. 2005. Virulence of oat crown rust in Brazil and Uruguay. Plant Dis. 89:802–808.
Lewis, C. M., Persoons, A., Bebber, D. P., Kigathi, R. N., Maintz, J., Findlay, K., et al. 2018. Potential for re-emergence of wheat stem rust in the United Kingdom. Commun. Biol. 1:13.
Li, F., Upadhyaya, N. M., Sperschneider, J., Matny, O., Nguyen-Phuc, H., Mago, R., et al. 2019. Emergence of the Ug99 lineage of the wheat stem rust pathogen through somatic hybridisation. Nat. Commun. 10:5068.
Li, H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 34:3094–3100.
Li, H., and Durbin, R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 25:1754–1760.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 25:2078–2079.
Lo Presti, L., Lanver, D., Schweizer, G., Reissman, S., and Kahmann, R. 2015. Fungal effectors and plant susceptibility. Annu. Rev. Plant Biol. 66:513–545.
Lorang, J. M., Sweat, T. A., and Wolpert, T. J. 2007. Plant disease susceptibility conferred by a “resistance” gene. Proc. Natl. Acad. Sci. 104:14861.
Lorrain, C., Gonçalves dos Santos, K. C., Germain, H., Hecker, A., and Duplessis, S. 2019. Advances in understanding obligate biotrophy in rust fungi. New Phytol. 222:1190–1206.
Lorrain, S., Lin, B., Auriac, M. C., Kroj, T., Saindrenan, P., Nicole, M., et al. 2004. Vascular associated death1, a novel GRAM domain-containing protein, is a regulator of cell death and defense responses in vascular tissues. Plant Cell. 16:2217–2232.
Love, M. I., Huber, W., and Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15:550.
Low, Y. C., Lawton, M. A., and Di, R. 2020. Validation of barley 2OGO gene as a functional orthologue of *Arabidopsis DMR6* gene in Fusarium head blight susceptibility. Sci. Rep. 10:9935.
Luig, N., and Watson, I. 1972. The role of wild and cultivated grasses in the

hybridization of formae speciales of *Puccinia graminis*. Aust. J. Biol. Sci. 25:335–342.
Luo, M., Li, H., Chakraborty, S., Morbitzer, R., Rinaldo, A., Upadhyaya, N., et al. 2019. Efficient TALEN-mediated gene editing in wheat. Plant Biotechnol. J. 17:2026–2028.
Manickavelu, A., Kawaura, K., Oishi, K., Shin-I, T., Kohara, Y., Yahiaoui, N., et al. 2010. Comparative gene expression analysis of susceptible and resistant near-isogenic lines in common wheat infected by *Puccinia triticina*. DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes. 17:211–222.
Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal. 17(1) doi:1014806ej171200.
Mayama, S., Bordin, A. P. A., Morikawa, T., Tanpo, H., and Kato, H. 1995. Association of avenalumin accumulation with co-segregation of victorin sensitivity and crown rust resistance in oat lines carrying the *Pc-2* gene. Physiol. Mol. Plant Pathol. 46:263–274.
McIntosh, R., Wellings, C., and Park, R. 1995. *Wheat Rusts: An atlas of resistance genes*. CSIRO.
Miller, M. E., Nazareno, E. S., Rottschaefer, S. M., Riddle, J., Dos Santos Pereira, D., Li, F., et al. 2021. Increased virulence of *Puccinia coronata* f. sp. *avenae* populations through allele frequency changes at multiple putative *Avr* loci. PLOS Genet. 16:e1009291.
Miller, M. E., Zhang, Y., Omidvar, V., Sperschneider, J., Schwessinger, B., Raley, C., et al. 2018. *De novo* assembly and phasing of dikaryotic genomes from two isolates of *Puccinia coronata* f. sp. *avenae*, the causal agent of oat crown rust mBio. 9:e01650-17.
Mitchell, A. G., and Martin, C. E. 1997. *Fah1p*, a *Saccharomyces cerevisiae* cytochrome$b_5$ fusion protein, and its *Arabidopsis thaliana* homolog that lacks the cytochrome$b_5$ domain both function in the α-hydroxylation of sphingolipid-associated very long chain fatty acids. J. Biol. Chem. 272:28281–28288.
Moerschbacher, B. M., Vander, P., Springer, C., Noll, U., and Schmittmann, G. 1994. Photosynthesis in stem rust-infected, resistant and susceptible near-isogenic wheat leaves. Can. J. Bot. 72:990–997.
Mundt, C. C. 2014. Durable resistance: A key to sustainable management of pathogens and pests. Infect. Genet. Evol. 27:446–455.
Murphy, H. C., and Meehan, F. 1946. Reaction of oat varieties to a new species of *Helminthosporium*. Phytopathology. 36:407.
Nagano, M., Takahara, K., Fujimoto, M., Tsutsumi, N., Uchimiya, H., and Kawai-Yamada, M. 2012. *Arabidopsis* sphingolipid fatty acid 2-hydroxylases (*AtFAH1* and *AtFAH2*) are functionally differentiated in fatty acid 2-hydroxylation and stress responses. Plant Physiol. 159:1138–1148.
Nazareno, E. S., Li, F., Smith, M., Park, R. F., Kianian, S. F., and Figueroa, M. 2018. *Puccinia coronata* f. sp. *avenae*: a threat to global oat production. Mol. Plant Pathol. 19:1047–1060.
Nekrasov, V., Wang, C., Win, J., Lanz, C., Weigel, D., and Kamoun, S. 2017. Rapid generation of a transgene-free powdery mildew resistant tomato by genome deletion. Sci. Rep. 7:482.

Nemri, A., Saunders, D., Anderson, C., Upadhyaya, N., Win, J., Lawrence, G., et al. 2014. The genome sequence and effector complement of the flax rust pathogen *Melampsora lini*. Front. Plant Sci. 5:98.

Olivera, P. D., Sikharulidze, Z., Dumbadze, R., Szabo, L. J., Newcomb, M., Natsarishvili, K., et al. 2019. Presence of a sexual population of *Puccinia graminis* f. sp. *tritici* in Georgia provides a hotspot for genotypic and phenotypic diversity. Phytopathology. 109:2152–2160.

Olivera, P., Newcomb, M., Szabo, L. J., Rouse, M., Johnson, J., Gale, S., et al. 2015. Phenotypic and genotypic characterization of race TKTTF of *Puccinia graminis* f. sp. *tritici* that caused a wheat stem rust epidemic in southern Ethiopia in 2013–14. Phytopathology. 105:917–928.

Omidvar, V., Dugyala, S., Li, F., Rottschaefer, S. M., Miller, M. E., Ayliffe, M., et al. 2018. Detection of race-specific resistance againt *Puccinia coronata* f. sp. *avenae* in *Brachypodium* species. Phytopathology. 108:1443–1454.

Ono, Y. 2002. The diversity of nuclear cycle in microcyclic rust fungi (Uredinales) and its ecological and evolutionary implications. Mycoscience. 43:421–439.

Palmer, J. M., and Stajich, J. 2020. *Funannotate v1.8.1: Eukaryotic genome annotation*. Zenodo. Available at: https://doi.org/10.5281/zenodo.4054262.

Paradis, E., and Schliep, K. 2018. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics. 35:526–528.

Periyannan, S., Milne, R. J., Figueroa, M., Lagudah, E. S., and Dodds, P. N. 2017. An overview of genetic rust resistance: from broad to specific mechanisms. PLoS Pathog. 13.

Pessina, S., Pavan, S., Catalano, D., Gallotta, A., Visser, R. G., Bai, Y., et al. 2014. Characterization of the *MLO* gene family in *Rosaceae* and gene expression analysis in *Malus domestica*. BMC Genomics. 15:618.

Petersen, R. H. 1974. The Rust Fungus Life Cycle. Bot. Rev. 40:453–513.

Peterson, P. D. 2001a. *Stem rust of wheat: from ancient enemy to modern foe*. St. Paul, USA. American Phytopathological Society.

Peterson, P. D. 2001b. The Campaign to Eradicate the Common Barberry in the United States. In *Stem Rust of Wheat: From Ancient Enemy to Modern Foe*, St. Paul, Minn.: Phytopathological Society.

Peterson, P. D. 2003. The common barberry: The past and present situation in Minnesota and the risk of wheat stem rust epidemics. Available at: https://www.ars.usda.gov/ARSUserFiles/50620500/Barberry/pdp_thesis.pdf.

Petit-Houdenot, Y., Langner, T., Harant, A., Win, J., and Kamoun, S. 2020. A clone resource of *Magnaporthe oryzae* effectors that share sequence and structural similarities across host-specific lineages. Mol. Plant-Microbe Interactions. 33:1032–1035.

Petre, B., Joly, D. L., and Duplessis, S. 2014. Effector proteins of rust fungi. Front. Plant Sci. 5:416.

Porto, B. N., Caixeta, E. T., Mathioni, S. M., Vidigal, P. M. P., Zambolim, L., Zambolim, E. M., et al. 2019. Genome sequencing and transcript analysis of *Hemileia vastatrix* reveal expression dynamics of candidate effectors dependent on host compatibility. PLOS ONE. 14:e0215598.

Pretorius, Z. A., Singh, R. P., Wagoire, W. W., and Payne, T. S. 2000. Detection of virulence to wheat stem rust resistance gene *Sr31* in *Puccinia graminis* f. sp. *tritici* in Uganda. Plant Dis. 84:203–203.

Robin, G. P., Kleemann, J., Neumann, U., Cabre, L., Dallery, J.-F., Lapalu, N., et al. 2018. Subcellular localization screening of *Colletotrichum higginsianum* effector candidates identifies fungal proteins targeted to plant peroxisomes, Golgi bodies, and microtubules. Front. Plant Sci. 9:562.

Roelfs, A. P. 1985. Wheat and Rye Stem Rust. In *The Cereal Rusts Volume II: Diseases, Distribution, Epidemiology, and Control*, Orlando, Florida: Academic Press, p. 3–37.

Rushton, P. J., Somssich, I. E., Ringler, P., and Shen, Q. J. 2010. WRKY transcription factors. Trends Plant Sci. 15:247–258.

Rutter, W. B., Salcedo, A., Akhunova, A., He, F., Wang, S., Liang, H., et al. 2017. Divergent and convergent modes of interaction between wheat and *Puccinia graminis* f. sp. *tritici* isolates revealed by the comparative gene co-expression network and genome analyses. BMC Genomics. 18:291.

Saintenac, C., Zhang, W., Salcedo, A., Rouse, M. N., Trick, H. N., Akhunov, E., et al. 2013. Identification of wheat gene *Sr35* that confers resistance to Ug99 stem rust race group. Science. 341:783.

Salcedo, A., Rutter, W., Wang, S., Akhunova, A., Bolus, S., Chao, S., et al. 2017. Variation in the *AvrSr35* gene determines *Sr35* resistance against wheat stem rust race Ug99. Science. 358:1604–1606.

Saur, I. M., Bauer, S., Kracher, B., Lu, X., Franzeskakis, L., Müller, M. C., et al. 2019. Multiple pairs of allelic *MLA* immune receptor-powdery mildew *AVRA* effectors argue for a direct recognition mechanism. eLife. 8:e44471.

van Schie, C. C., and Takken, F. L. 2014. Susceptibility genes 101: how to be a good host. Annu. Rev. Phytopathol. 52:551–581.

Schwessinger, B., Chen, Y.-J., Tien, R., Vogt, J. K., Sperschneider, J., Nagar, R., et al. 2020. Distinct life histories impact dikaryotic genome evolution in the rust fungus *Puccinia striiformis* causing stripe rust in wheat. Genome Biol. Evol. 12:597–617.

Schwessinger, B., Sperschneider, J., Cuddy, W. S., Garnica, D. P., Miller, M. E., Taylor, J. M., et al. 2018. A near-complete haplotype-phased genome of the dikaryotic wheat stripe rust fungus *Puccinia striiformis* f. sp. *tritici* reveals high interhaplotype diversity. mBio. 9:e02275-17.

Servant, N., Varoquaux, N., Lajoie, B. R., Viara, E., Chen, C.-J., Vert, J.-P., et al. 2015. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol. 16:259.

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

Simons, M. D. 1985. Crown Rust. In *The Cereal Rusts Volume II: Diseases, Distribution, Epidemiology, and Control*, Orlando, Florida, p. 131–172.

Singh, R. P., Hodson, D. P., Huerta-Espino, J., Jin, Y., Bhavani, S., Njau, P., et al. 2011. The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. Annu. Rev. Phytopathol. 49:465–481.

Singh, R. P., Hodson, D. P., Jin, Y., Lagudah, E. S., Ayliffe, M. A., Bhavani, S., et al. 2015. Emergence and spread of new races of wheat stem rust fungus: continued threat to food security and prospects of genetic control. Phytopathology. 105:872–884.

Smit, A. F. A., Hubley, R., and Green, P. 2013. *RepeatMasker Open-4.0*. Available at: http://www.repeatmasker.org.

Song, J. T., Lu, H., and Greenberg, J. T. 2004. Divergent roles in *Arabidopsis thaliana* development and defense of two homologous genes, aberrant growth and death2 and AGD2-LIKE DEFENSE RESPONSE PROTEIN1, encoding novel aminotransferases. Plant Cell. 16:353–366.

Staples, R., and Macko, V. 1984. Germination of urediospores and differentiation of infection structures. In *The Cereal Rusts*, ed. William Bushnell. Academic Press, p. 255–289.

Steffenson, B. J., Case, A. J., Pretorius, Z. A., Coetzee, V., Kloppers, F. J., Zhou, H., et al. 2017. Vulnerability of barley to African pathotypes of *Puccinia graminis* f. sp. *tritici* and sources of resistance. Phytopathology. 107:950–962.

Sun, K., van Tuinen, A., van Kan, J. A. L., Wolters, A.-M. A., Jacobsen, E., Visser, R. G. F., et al. 2017. Silencing of *DND1* in potato and tomato impedes conidial germination, attachment and hyphal growth of *Botrytis cinerea*. BMC Plant Biol. 17:235.

Sun, K., Wolters, A.-M. A., Vossen, J. H., Rouwet, M. E., Loonen, A. E. H. M., Jacobsen, E., et al. 2016. Silencing of six susceptibility genes results in potato late blight resistance. Transgenic Res. 25:731–742.

Tavares, S., Ramos, A. P., Pires, A. S., Azinheira, H. G., Caldeirinha, P., Link, T., et al. 2014. Genome size analyses of Pucciniales reveal the largest fungal genomes. Front. Plant Sci. 5:422.

Thorvaldsdóttir, H., Robinson, J. T., and Mesirov, J. P. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. 14:178–192.

Upadhyaya, N. M., Garnica, D. P., Karaoglu, H., Sperschneider, J., Nemri, A., Xu, B., et al. 2015. Comparative genomics of Australian isolates of the wheat stem rust pathogen *Puccinia graminis* f. sp. *tritici* reveals extensive polymorphism in candidate effector genes. Front. Plant Sci. 5:759.

Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S., and Gregory, P. D. 2010. Genome editing with engineered zinc finger nucleases. Nat. Rev. Genet. 11:636–646.

USDA-ARS CDL. 2014. 2014 Oat Loss to Rust. Available at: https://www.ars.usda.gov/ARSUserFiles/50620500/Smallgrainlossesduetorust/2014loss/2014oatloss.pdf.

Van Damme, M., Huibers, R. P., Elberse, J., and Van den Ackerveken, G. 2008. *Arabidopsis DMR6* encodes a putative 2OG-Fe(II) oxygenase that is defense-associated but required for susceptibility to downy mildew. Plant J. 54:785–793.

Vogel, J., and Hill, T. 2008. High-efficiency *Agrobacterium*-mediated transformation of *Brachypodium distachyon* inbred line Bd21-3. Plant Cell Rep. 27:471–478.

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. 2014. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLOS ONE. 9:e112963.

Wang, W., and Jiao, F. 2019. Effectors of *Phytophthora* pathogens are powerful weapons for manipulating host immunity. Planta. 250:413–425.
Wang, X., Tang, C., Huang, X., Li, F., Chen, X., Zhang, G., et al. 2012. Wheat BAX inhibitor-1 contributes to wheat resistance to *Puccinia striiformis*. J. Exp. Bot. 63:4571–4584.
Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C., et al. 2014. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. Nat. Biotechnol. 32:947–951.
Waterhouse, P. M., and Helliwell, C. A. 2003. Exploring plant genomes by RNA-induced gene silencing. Nat. Rev. Genet. 4:29–38.
Welsh, J. N., Peturson, B., and Machacek, J. E. 1954. Associated inheritance of reaction to races of crown rust, *Puccinia coronata avenae* Erikss., and to Victoria blight, *Helminthosporium victoriae* M. and M., in oats. Can. J. Bot. 32:55–68.
Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.
Wildermuth, M. C. 2010. Modulation of host nuclear ploidy: a common plant biotroph mechanism. Curr. Opin. Plant Biol. 13:449–458.
Williams, P. G. 1984. Obligate parasitism and axenic culture. In *The Cereal Rusts: Origins, Specificity, Structure, and Physiology*, Orlando, Florida: Academic Press.
Wolff, J., Rabbani, L., Gilsbach, R., Richard, G., Manke, T., Backofen, R., et al. 2020. Galaxy HiCExplorer 3: a web server for reproducible Hi-C, capture Hi-C and single-cell Hi-C data analysis, quality control and visualization. Nucleic Acids Res. 48:W177–W184.
Wolpert, T. J., and Macko, V. 1989. Specific binding of victorin to a 100-kDa protein from oats. Proc. Natl. Acad. Sci. U. S. A. 86:4092–4096.
Wolpert, T. J., Macko, V., Acklin, W., Jaun, B., Seibl, J., Meili, J., et al. 1985. Structure of victorin C, the major host-selective toxin from *Cochliobolus victoriae*. Experientia. 41:1524–1529.
Wu, J. Q., Dong, C., Song, L., and Park, R. F. 2020. Long-read–based *de novo* genome assembly and comparative genomics of the wheat leaf rust pathogen *Puccinia triticina* identifies candidates for three avirulence genes. Front. Genet. 11:521.
Xia, C., Wang, M., Yin, C., Cornejo, O. E., Hulbert, S. H., and Chen, X. 2018. Genome sequence resources for the wheat stripe rust pathogen (*Puccinia striiformis* f. sp. *tritici*) and the barley stripe rust pathogen (*Puccinia striiformis* f. sp. *hordei*). Mol. Plant-Microbe Interactions. 31:1117–1120.
Yadav, I. S., Sharma, A., Kaur, S., Nahar, N., Bhardwaj, S. C., Sharma, T. R., et al. 2016. Comparative temporal transcriptome profiling of wheat near isogenic line carrying *Lr57* under compatible and incompatible interactions. Front. Plant Sci. 7:1943.
Yu, G., Smith, D. K., Zhu, H., Guan, Y., and Lam, T. T.-Y. 2017. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. 8:28–36.
Yu, I., Fengler, K. A., Clough, S. J., and Bent, A. F. 2000. Identification of *Arabidopsis* mutants exhibiting an altered hypersensitive response in gene-for-gene disease resistance. Mol. Plant-Microbe Interactions. 13:277–286.

Zaidi, S. S.-A., Mukhtar, M. S., and Mansoor, S. 2018. Genome editing: Targeting susceptibility genes for plant disease resistance. Trends Biotechnol. 36:898–906.

Zambino, P. J., Kubelik, A. R., and Szabo, L. J. 2000. Gene action and linkage of avirulence genes to DNA markers in the rust fungus *Puccinia graminis*. Phytopathology. 90:819–826.

Zeilmaker, T., Ludwig, N. R., Elberse, J., Seidl, M. F., Berke, L., Van Doorn, A., et al. 2015. DOWNY MILDEW RESISTANT 6 and DMR6-LIKE OXYGENASE 1 are partially redundant but distinct suppressors of immunity in *Arabidopsis*. Plant J. 81:210–222.

Zhang, D., Bowden, R. L., Yu, J., Carver, B. F., and Bai, G. 2014. Association analysis of stem rust resistance in U.S. winter wheat. PLOS ONE. 9:e103747.

Zhang, Y., Zhao, L., Zhao, J., Li, Y., Wang, J., Guo, R., et al. 2017. *S5H/DMR6* encodes a salicylic acid 5-hydroxylase that fine-tunes salicylic acid homeostasis. Plant Physiol. 175:1082–1093.

Zheng, W., Huang, L., Huang, J., Wang, X., Chen, X., Zhao, J., et al. 2013. High genome heterozygosity and endemic genetic recombination in the wheat stripe rust fungus. Nat. Commun. 4:2673.

Zheng, Z., Mosher, S. L., Fan, B., Klessig, D. F., and Chen, Z. 2007. Functional analysis of *Arabidopsis WRKY25* transcription factor in plant defense against *Pseudomonas syringae*. BMC Plant Biol. 7:2.

# Appendices

## Appendix A: Emergence of the Ug99 lineage of the wheat stem rust pathogen through somatic hybridization[2]



**Abstract**

Parasexuality contributes to diversity and adaptive evolution of haploid (monokaryotic) fungi. However, non-sexual genetic exchange mechanisms are not defined in dikaryotic fungi (containing two distinct haploid nuclei). Newly emerged strains of the wheat stem rust pathogen, *Puccinia graminis* f. sp. *tritici* (*Pgt*), such as Ug99, are a major threat to global food security. Here, we provide genomics-based evidence supporting that Ug99 arose by somatic hybridisation and nuclear exchange between dikaryons. Fully haplotype-resolved genome assembly and DNA proximity analysis reveal that Ug99 shares one haploid nucleus genotype with a much older African lineage of *Pgt*, with no recombination or chromosome reassortment. These findings indicate that nuclear exchange between dikaryotes can generate genetic diversity and facilitate the emergence of new lineages in asexual fungal populations.

---

[2] Li, F., Upadhyaya, N.M., Sperschneider, J. *et al.* 2019. Emergence of the Ug99 lineage of the wheat stem rust pathogen through somatic hybridisation. *Nat Commun.*10:5068

**Contribution to the work:**

Responsible for orthology analyses during the haplotype assignment steps and manuscript revisions

# Appendix B: Increased virulence of *Puccinia coronata* f. sp. *avenae* populations through allele frequency changes at multiple putative Avr loci[3]



**Abstract**
Pathogen populations are expected to evolve virulence traits in response to resistance deployed in agricultural settings. However, few temporal datasets have been available to characterize this process at the population level. Here, we examined two temporally separated populations of *Puccinia coronata* f. sp. *avenae* (*Pca*), which causes crown rust disease in oat (*Avena sativa*) sampled from 1990 to 2015. We show that a substantial increase in virulence occurred from 1990 to 2015 and this was associated with a genetic differentiation between populations detected by genome-wide sequencing. We found strong evidence for genetic recombination in these populations, showing the importance of the alternate host in generating genotypic variation through sexual reproduction. However, asexual expansion of some clonal lineages was also observed within years. Genome-wide association analysis identified seven *Avr* loci associated with virulence towards fifteen *Pc* resistance genes in oat and suggests that some groups of *Pc* genes recognize the same pathogen effectors. The temporal shift in virulence patterns in the *Pca* populations between 1990 and 2015 is associated with changes in allele frequency in these genomic regions. Nucleotide diversity patterns at a single *Avr* locus corresponding to *Pc38*, *Pc39*, *Pc55*, *Pc63*, *Pc70*, and *Pc71* showed evidence of a selective sweep associated with the shift to virulence towards these resistance genes in all 2015 collected isolates.

---

[3] Miller M.E., Nazareno E.S., Rottschaefer S.M., Riddle J., Dos Santos Pereira D., et al. 2020. Increased virulence of *Puccinia coronata f. sp. avenae* populations through allele frequency changes at multiple putative *Avr* loci. PLOS Genetics. 16(12): e1009291.

**Contribution to this work:**
Involved with the interpretation of GWAS results and manuscript revisions.

# Appendix C: Evolution of virulence in rust fungi — multiple solutions to one problem[4]



**Abstract**
Rust fungi are major pathogens that negatively affect crops and ecosystems. Recent rust disease epidemics driven by the emergence of strains with novel virulence profiles demand a better understanding of the evolutionary mechanisms of these organisms. Here, we review research advances in genome-scale analysis coupled with functional validation of effector candidate genes that have been instrumental to elucidate processes that contribute to changes in virulence phenotypes. We highlight how haplotype-phased genome references have paved the road to link these processes to the reproductive phases of rust fungi and have provided evidence for somatic exchange between strains as an important mechanism for generating diversity in asexual populations. With increasing data availability, we envision the future development of molecular virulence diagnostic tools.

---

[4] Figueroa, M., Dodds, P. N., and Henningsen, E. C. 2020. Evolution of virulence in rust fungi — multiple solutions to one problem. Curr. Opin. Plant Biol. 56:20–27

**Contribution to this work:**
Contributed to writing, editing, and revising the manuscript draft.

# Appendix D: *Rpg7*: A New Gene for Stem Rust Resistance from *Hordeum vulgare* ssp. *spontaneum*[5]



**Abstract**

Wheat stem rust (causal organism: *Puccinia graminis* f. sp. *tritici*) is an important fungal disease that causes significant yield losses in barley. The deployment of resistant cultivars is the most effective means of controlling this disease. Stem rust evaluations of a diverse collection of wild barley (*Hordeum vulgare* ssp. *spontaneum*) identified two Jordanian accessions (WBDC094 and WBDC238) with resistance to a virulent pathotype (*P. graminis* f. sp. *tritici* HKHJC) from the United States. To elucidate the genetics of stem rust resistance, both accessions were crossed to the susceptible landrace Hiproly. Segregation ratios of $F_2$ and $F_3$ progeny indicated that a single dominant gene confers resistance to *P. graminis* f. sp. *tritici* HKHJC. Molecular mapping of the resistance locus was performed in the Hiproly/WBDC238 $F_2$ population based on 3,329 single-nucleotide polymorphism markers generated by genotyping-by-sequencing. Quantitative trait locus analysis positioned the resistance gene to the long arm of chromosome 3H between the physical/genetic positions of 683.8 Mbp/172.9 cM and 693.7 Mbp/176.0 cM. Because this resistance gene is novel, it was assigned the new gene locus symbol of *Rpg7* with a corresponding allele symbol of *Rpg7.i*. At the seedling stage, *Rpg7* confers resistance against a number of other important *P. graminis* f. sp. *tritici* pathotypes from the United States (MCCFC, QCCJB, and TTTTF) and Africa (TTKSK) as well as an isolate (92-MN-90) of the rye stem rust pathogen (*P. graminis* f. sp. *secalis*) from Minnesota. The resistance conferred by *Rpg7* can be readily transferred into breeding programs because of its simple inheritance and clear phenotypic expression.

---

[5] Henningsen, E.C., Sallam, A.H., Matny, O., Szinyei, T., Figueroa, M., Steffenson, B.J. 2021. *Rpg7*: A New Gene for Stem Rust Resistance from *Hordeum vulgare* ssp. *spontaneum*. Phytopathology. 111(3):548-558

**Contribution to this work:**

I conducted the inoculation experiments, scoring, and managed data entry for $F_2$, $F_3$, and allelism experiments. I also filtered GBS marker data, created linkage and QTL maps, and generated plots, and identified candidate genes from the QTL results. I also wrote, edited, submitted, and revised the manuscript.

# Appendix E: Figueroa et al. Tactics of host manipulation by intracellular effectors from plant pathogenic fungi[6]



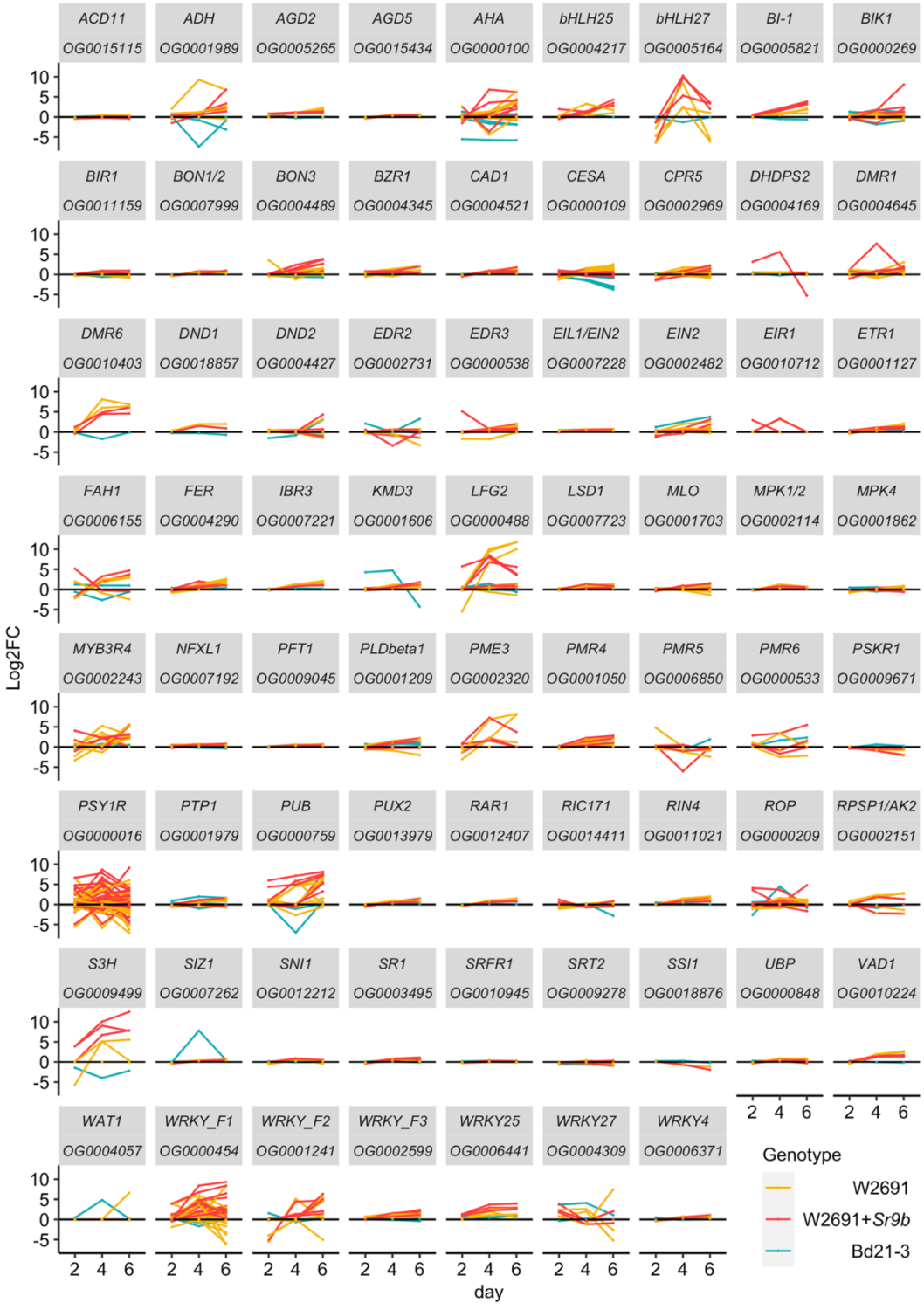
**Abstract**
Fungal pathogens can secrete hundreds of effectors, some of which are known to promote host susceptibility. This biological complexity, together with the lack of genetic tools in some fungi, presents a substantial challenge to develop a broad picture of the mechanisms these pathogens use for host manipulation. Nevertheless, recent advances in understanding individual effector functions are beginning to flesh out our view of fungal pathogenesis. This review discusses some of the latest findings that illustrate how effectors from diverse species use similar strategies to modulate plant physiology to their advantage. We also summarize recent breakthroughs in the identification of effectors from challenging systems, like obligate biotrophs, and emerging concepts such as the "iceberg model" to explain how the activation of plant immunity can be turned off by effectors with suppressive activity.

---

[6] Figueroa, M. Ortiz, D. Henningsen, E.C. 2021. Tactics of host manipulation by intracellular effectors from plant pathogenic fungi. Current Opinions in Plant Biology. 62

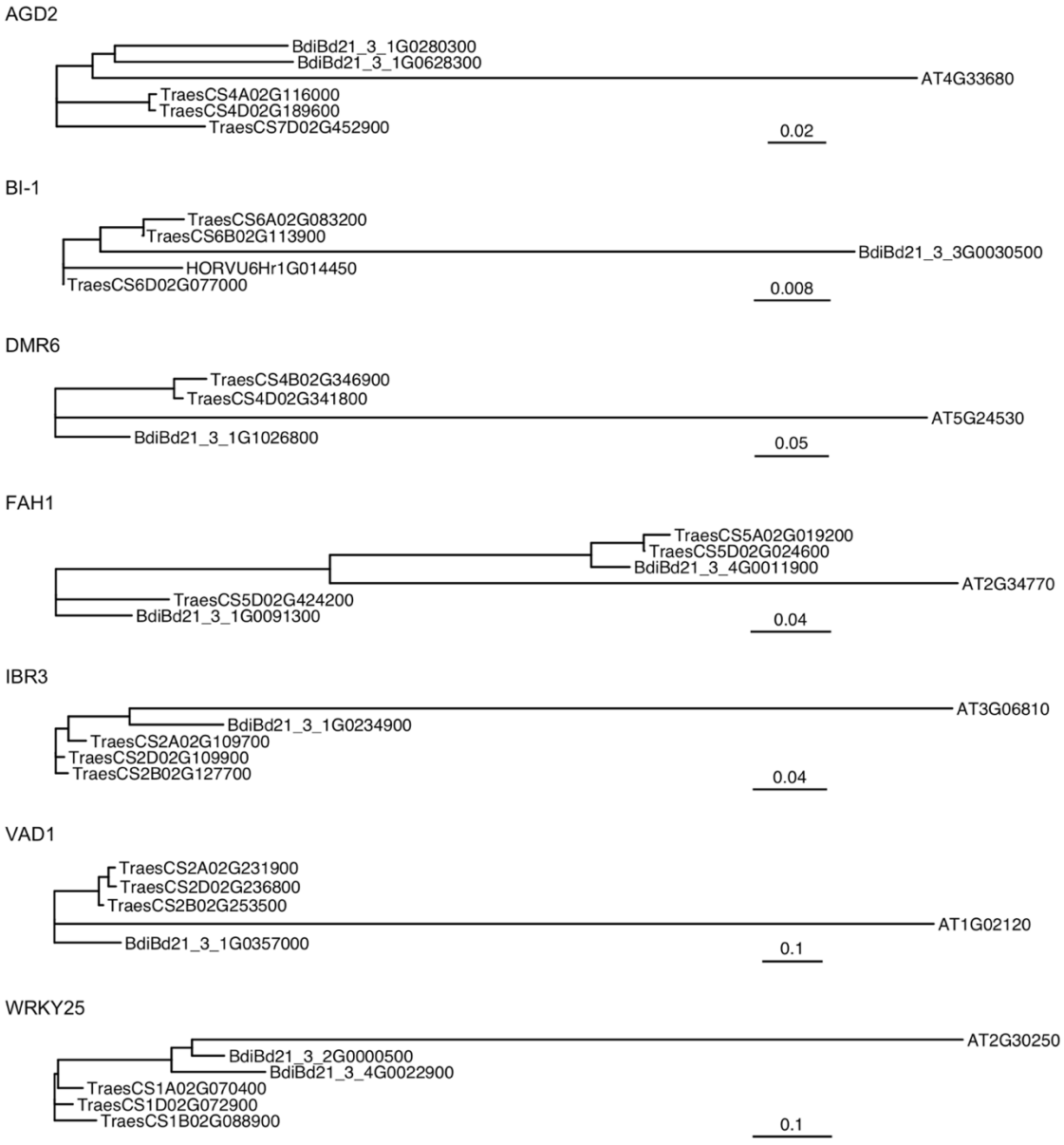
**Contribution to this work:**
Contributed to the writing, editing, and revision of the manuscript draft.

# Appendix F: Supplementary Materials for Chapter 2

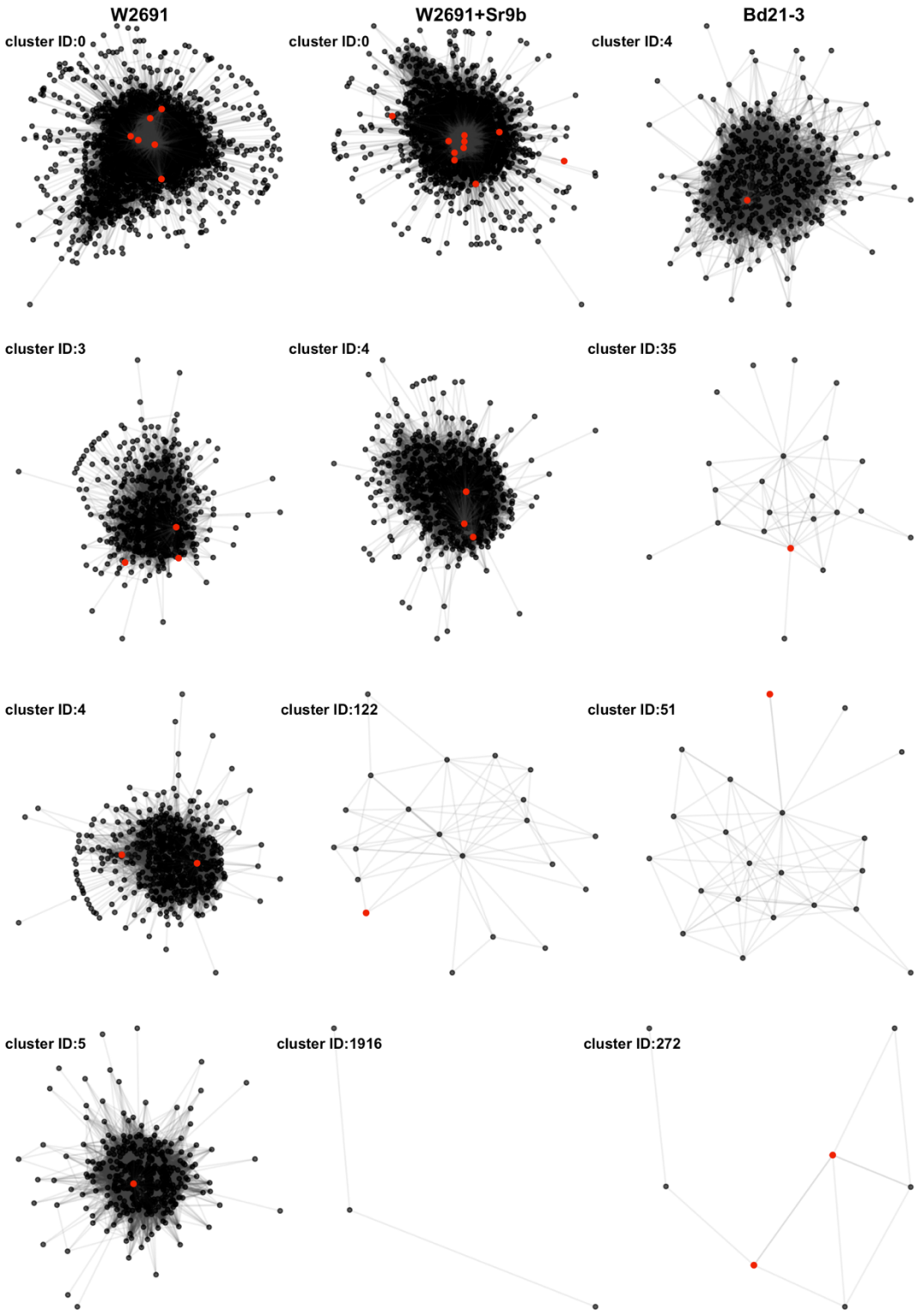


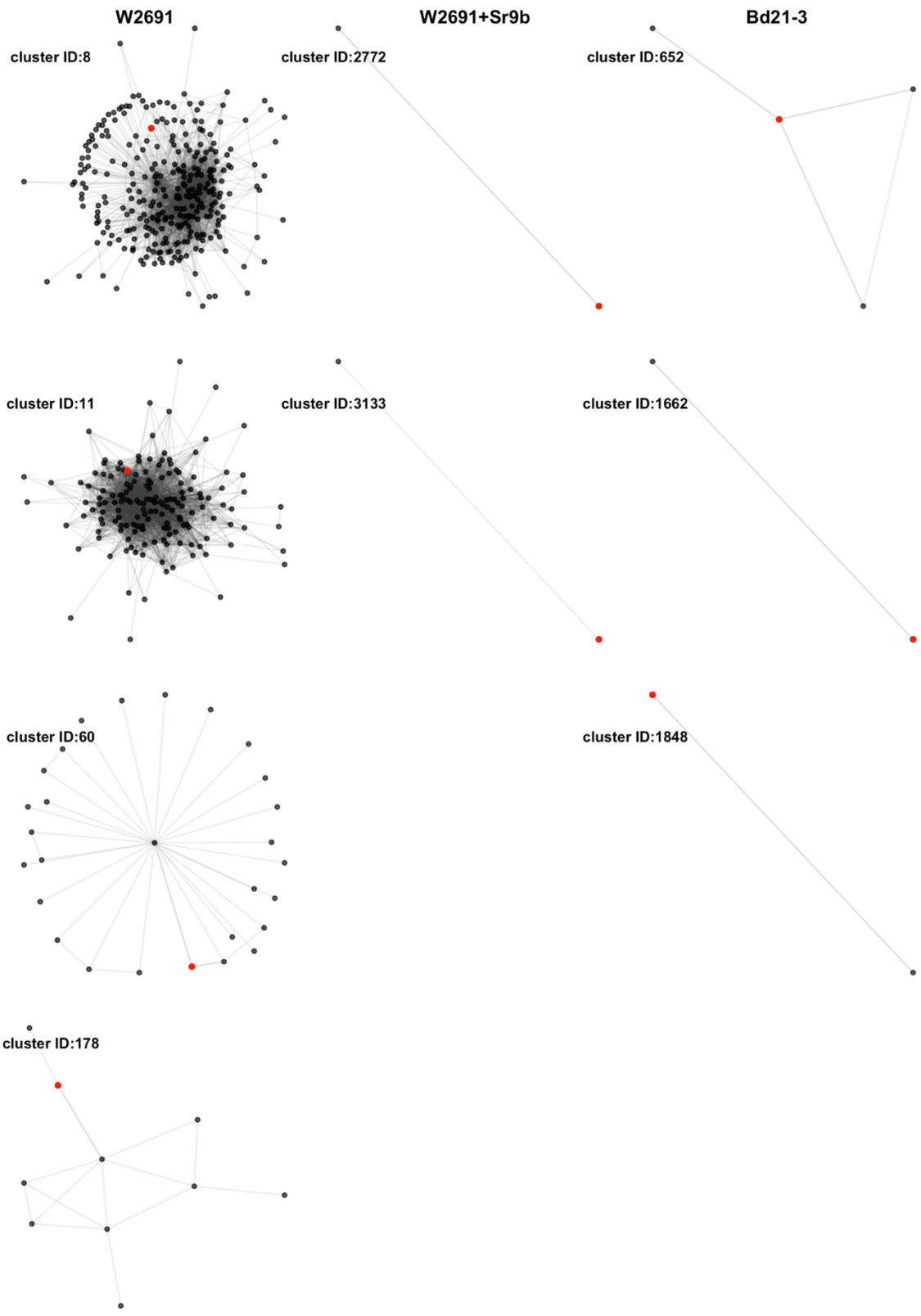

**Supplemental Figure 1.** Expression profile patterns of 80 orthogroups containing candidate *S* genes in *T. aestivum* ('W2691' and 'W2691+*Sr9b*') and *B. distachyon* (Bd21-

3) genotypes throughout infection with *P. graminis* f. sp. *tritici*. Log2 fold change values (y-axis) for all gene orthologs are presented per sampling time point (x-axis). Name of *S* gene and orthogroup identifier are shown in each graph.




**Supplemental Figure 2.** Molecular phylogenetic analysis of amino acid sequences of orthologous genes for five of the six S genes of interest. The orthogroup for DND1 only included three genes (One *A. thaliana* susceptibility gene, one *T. aestivum* ortholog, and one *B. distachyon* ortholog) and so no phylogenetic tree was generated. Gene names highlighted in red are those represented in Figure 3. Scale bars represent nucleotide substitutions per site.


W2691
W2691+Sr9b
Bd21-3
cluster ID:0
cluster ID:0
cluster ID:4
cluster ID:3
cluster ID:4
cluster ID:35
cluster ID:4
cluster ID:122
cluster ID:51
cluster ID:5
cluster ID:1916
cluster ID:272


W2691
W2691+Sr9b
Bd21-3
cluster ID:8
cluster ID:2772
cluster ID:652
cluster ID:11
cluster ID:3133
cluster ID:1662
cluster ID:60
cluster ID:1848
cluster ID:178

**Supplemental Figure 3.** All clusters (nodes > 1) containing a *T. aestivum* or *Brachypodium distachyon* ortholog of the eight susceptibility candidates. Red points counts represent *T. aestivum* and *B. distachyon* orthologs of *S* genes. Gene IDs of members in coexpression clusters are presented in Table S9.

# Appendix G: Supplementary Materials for Chapter 3

**Supplemental Table 1**. Summary of contigs that were excluded from the final assembly.

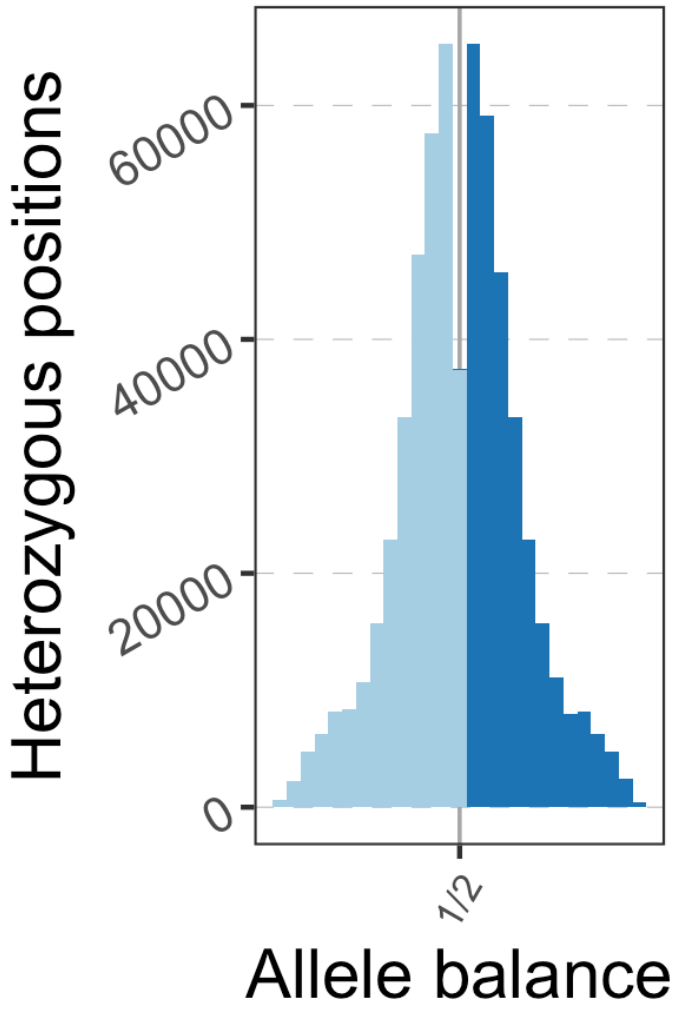
| Contig name | Reason for exclusion | Length (bp) | Coverage |
|---|---|---|---|
| tig397 | low coverage | 36920 | 0.3228 |
| tig492 | oat contamination | 13331 | 100.8359 |
| tig493 | oat contamination | 18720 | 20.5608 |
| tig494 | mitochondrial genome | 2523 | 84.7832 |
| tig496 | PacBio construct | 3155 | 2284.497 |
| tig504 | sequencing error | 72066 | 0* |
| tig505 | bacterial contamination | 2867 | 19.4513 |
| tig506 | insect contamination | 2841 | 134.868 |
| tig507 | Brachypodium contamination | 8353 | 59.8489 |
| tig523 | low coverage | 29618 | 1.4003 |
| tig718 | low coverage | 27695 | 1.3597 |
| tig719 | low coverage | 27660 | 1.229 |
| tig726 | low coverage | 39187 | 1.8858 |
| tig739 | low coverage | 25489 | 1.649 |
| tig777 | low coverage | 28198 | 1.5195 |
| tig848 | mitochondrial genome | 115256 | 14557.4051 |
| tig849 | mitochondrial genome | 64154 | 1126.7427 |
| tig852 | oat contamination | 6773 | 7.0171 |
| tig853 | oat contamination | 13110 | 34.335 |

**Supplemental Table 2**. Summary of bin alignments to the *Puccinia graminis* f. sp. *tritici* isolate 21-0 A haplotype.

| Pgt21-0 Chromosome | # of assigned 203 bins | 203 bin identifiers | Length of Pgt21-0 chromosome |
|---|---|---|---|
| 1A | 3 | 13, 24, 44 | 6156719 |
| 2A | 4 | 12, 14, 28, 43 | 6062785 |
| 3A | 2 | 5, 17 | 6034716 |
| 4A | 4 | 9, 27, 38, 40 | 5967513 |
| 5A | 3 | 15, 18, 23 | 5558200 |
| 6A | 1 | 10 | 5554679 |
| 7A | 3 | 1, 21, 39 | 5184328 |
| 8A | 1 | 8 | 5113503 |
| 9A | 4 | 4, 31, 36, 45 | 4787822 |
| 10A | 3 | 3, 33, 41 | 4648557 |
| 11A | 2 | 6, 22 | 4640345 |
| 12A | 2 | 25, 34 | 3976801 |
| 13A | 3 | 16, 30, 37 | 3570070 |
| 14A | 2 | 11, 26 | 3568213 |
| 15A | 1 | 20, 32 | 3494038 |
| 16A | 1 | 7 | 3430719 |
| 17A | 3 | 19, 29, 35 | 3318234 |
| 18A | 1 | 2 | 2873395 |

| Total length of assigned bins haplotype A | Total length of bins haplotype B | % length covered haplotype A |
|---|---|---|
| 6944168 | 7401722 | 112.79 |
| 6399800 | 6672529 | 105.56 |
| 6326871 | 6653677 | 104.84 |
| 7148849 | 8085236 | 119.80 |
| 6848375 | 6685610 | 123.21 |
| 5027340 | 5135709 | 90.51 |
| 5316925 | 6192698 | 102.56 |
| 5162947 | 5257689 | 100.97 |
| 3569350 | 3348622 | 74.55 |
| 3756089 | 4113161 | 80.80 |
| 5515009 | 5570172 | 118.85 |
| 4005011 | 4683873 | 100.71 |
| 5191522 | 5084385 | 145.42 |
| 2974161 | 3507466 | 83.35 |
| 3255862 | 3293895 | 93.18 |
| 3849125 | 4353638 | 112.20 |
| 3852422 | 3963856 | 116.10 |
| 2931919 | 3412090 | 102.04 |

| % length covered haplotype B |
|---|
| 120.22 |
| 110.06 |
| 110.26 |
| 135.49 |
| 120.28 |
| 92.46 |
| 119.45 |
| 102.82 |
| 69.94 |
| 88.48 |
| 120.04 |
| 117.78 |
| 142.42 |
| 98.30 |
| 94.27 |
| 126.90 |
| 119.46 |
| 118.75 |




**Supplemental Figure 1.** Allele balance plot of variants for *Pca* 203 as compared to the existing 12SD80 reference genome.