Whole Genome Sequencing and Comparative
Genomics of Fusarium Head Blight Fungi

by

Sean Walkowiak

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Biology

Carleton University
Ottawa, Ontario

© 2016

Sean Walkowiak

ABSTRACT

Plant pathogenic fungi from the genus *Fusarium* can cause economically important diseases of several different agricultural crops, including Fusarium Head Blight (FHB) of wheat. In addition to physically damaging the crops and causing reduced yield, these fungi produce toxins that contaminate the grains and make their products unsafe for consumption. Fungi causing this disease are considerably diverse in their biology with the most prominent and globally distributed species being from the *F. graminearum* species complex. Fungi from other taxonomic groups, such as *F. avenaceum* from the *F. tricinctum* species complex, also cause FHB and have the same geographical ranges as FHB pathogens from the *F. graminearum* species complex. This study explores the genomic similarities and differences between FHB fungi to: 1) better understand their diversity, and 2) identify genes that may be contributing to toxin production and disease. Whole genome sequencing and comparative genomics was performed on three species within the *F. graminearum* species complex; *F. graminearum*, *F. meridionale*, and *F. asiaticum*. Patterns of genomic and genetic variation were used to identify genomic regions and genes that are different between genomes. We then performed more detailed analyses of the biological interactions between two North American strains of *F. graminearum* that are from distinctly different genetic populations and produce different toxins. Our analyses indicated that one strain is more competitive than the other and that isolate-isolate interactions negatively impact toxin production and disease. We also identified a region of the genome that is highly variable between the two strains and when this region was compared to other strains of *F. graminearum*, the analyses indicated it is part of the accessory genome of the *F. graminearum* species and genus. Finally, we also performed whole genome sequencing and comparative genomics of three *F. avenaceum* isolates, which are FHB pathogens within the *F.*

tricinctum species complex that are distantly related to the *F. graminearum* species complex.

Our analyses indicated that the *F. avenaceum* genome is larger than that of *F. graminearum* and contains a different genetic repertoire that may contribute to toxin production and disease.

ACKNOWLEDGEMENTS

The research presented in this thesis encompasses a large body of work, with contributions from others, either in the form of technical support on the various research projects or as contributions to my personal and professional development and well-being, both of which were instrumental to the success of my research. There have been countless people who have supported me up until this specific point in time, and I am truly grateful to all of them for the important roles that they have had.

To begin, I thank Drs. Rajagopal (Gopal) Subramaniam and Owen Rowland for their support as graduate supervisors. They have both been active mentors and have instilled the passion for science that I now carry with me. I first met Owen as an undergraduate student enrolled in his courses at Carleton University, where Owen inspired me to apply myself as an inventive scientist. I was then introduced to Gopal, who welcomed me into his research laboratory at the Ottawa Research and Development Center (ORDC) at Agriculture and Agri-Food Canada (AAFC) and mentored me on how to be a scientific researcher; he taught me everything from basic laboratory skills to the importance of developing and testing innovative research questions. Both Owen and Gopal have supported and guided me throughout my experiences as an undergraduate and graduate student, and have been thoroughly engaged in my development as a scientist, for which I will be forever grateful.

I recognize the support of other members of my immediate research team at AAFC ORDC who guided me and my research and provided technical assistance and advice (appearing alphabetically): Christopher Bonner, Elizabeth Brauer, Salima Chatur, Margaret Crew, Manisha Joshi, Winnie Leung, Samantha Lewis, Armand Mirmiran, Christopher Mogg, Swara Narayanan,

Charles Nasmith, Kristina Shostak, Li Wang, and Adamo Young. Although all members of the research team have assisted me throughout the years, one person in particular has guided me through my entire journey in research, and together with Dr. Subramaniam has been the foundation of the research laboratory; this person is Li Wang. Li is a very warm and welcoming person who has never hesitated to offer support, whether it was technical support or just a listening ear, Li was always there.

Outside of our immediate research team, there have been a number of collaborators that have contributed to my research or professional development. One of the most satisfying experiences I have had in research has been interacting with other scientists, sharing research ideas, and working together on different research projects. I acknowledge and appreciate Drs. Linda Harris and Erik Lysøe for allowing me to contribute to their project on the sequencing of the *Fusarium avenaceum* genome. Dr. Harris and her team also provided technical support for digital PCR and microarrays as part of other projects associated with my research. I thank Dr. Christof Rampitsch, for sharing his research ideas and expertise on protein mass spectrometry on various collaborative projects we have worked on together. I recognize Dr. Barbara Blackwell and her research lab for their analytical chemistry expertise that they have provided. I also acknowledge the members of my research and evaluation committees, past and present, for providing guidance during my research: Drs. Thérèse Ouellet, Alex Wong, Nicolas Rodrigue, and Myron Smith.

Finally, and perhaps the greatest supporters of my research, I acknowledge the support of my family. My wife Rebecca, our two children, Allison and Elizabeth, and my parents Peter and Marion have encouraged and supported me throughout my scientific career.

PREFACE

Each data chapter from this thesis are adapted from research manuscripts that have been published, submitted, or are in preparation for submission in which S.W is a primary or co-author. The detailed reference for each manuscript and author contributions for each chapter are provided.

Chapter 2:

This chapter is adapted from: Walkowiak , S., O. Rowland and R. Subramaniam. (Submitted).

"Whole genome sequencing and comparative genomics of closely related Fusarium Head Blight fungi: *Fusarium graminearum*, *F. meridionale* and *F. asiaticum*" BMC Genomics. S.W is the main contributor to this manuscript and performed all data analyses and interpretations; R.S and O.R contributed to manuscript writing and editing.

Chapter 3:

This chapter is adapted from: Walkowiak, S., C. T. Bonner, L. Wang, B. Blackwell, O. Rowland and R. Subramaniam (2015). "Intraspecies interaction of *Fusarium graminearum* contributes to reduced toxin production and virulence." Molecular Plant-Microbe Interactions 28(11): 1256-1267. S.W is the primary contributor to the manuscript, and performed all experiments included in the thesis with the exception of some fungal culturing, high performance liquid chromatography (HPLC) analyses, and polymerase chain reaction (PCR) experiments, which were performed by CT.B. Also, L.W performed DNA isolation and B.B performed nuclear magnetic resonance (NMR) of secondary metabolites. R.S and O.R contributed to manuscript writing and editing.

Chapter 4:

This chapter is a manuscript in preparation with the main contributors being Mirmiran, A.*, S. Walkowiak*, M. Crew, O. Rowland, A. Johnston and R. Subramaniam. The ‘*’ indicates equal contribution and co-primary authorship of the manuscript by A.M and S.W. Contributions to the chapter by S.W include the experimental design and data interpretations, the discovery of contig_207 and its variation amongst genomes, *in silico* investigations of contig structure and gene functions. A.M performed all other experiments, with assistance from M.C, under the supervision and guidance of S.W. Figures 4.1 and 4.5 were prepared cooperatively by A.M and S.W, all other figures were prepared by S.W. A.J. conducted the digital PCR analyses. Assistance with experimental design and data interpretations was provided by O.R and R.S.

Chapter 5:

This chapter is adapted from: Lysøe, E., L. J. Harris, S. Walkowiak, R. Subramaniam, H. H. Divon, E. S. Riiser, C. Llorens, T. Gabaldón, H. C. Kistler, W. Jonkers, A.-K. Kolseth, K. F. Nielsen, U. Thrane and R. J. N. Frandsen (2014). "The genome of the generalist plant pathogen *Fusarium avenaceum* is enriched with genes involved in redox, signaling and secondary metabolism." PLoS ONE 9(11): doi: 10.1371/journal.pone.0112703. Only a subset of the manuscript that included contributions from S.W has been included in the thesis. Contributions to the manuscript by S.W include the *de novo* genome assemblies and gene prediction for the two Canadian *F. avenaceum* isolates, the higher order assembly and comparative genomics of all three *F. avenaceum* isolates, the genome alignments of *F. avenaceum* to each other and to other fusaria, and the comparisons in gene content within *F. avenaceum* and to members of the *F. graminearum* species complex by BLASTn.

TABLE OF CONTENTS

Abstract	i
Acknowledgements	iii
Preface	v
Table of Contents	vii
List of Tables	xii
List of Figures	xiii
List of Abbreviations	xvi

Chapter 1: Introduction

1.1 Fusarium Head Blight (FHB): Impacts to the Economy and Human Health ..	1
1.2 The Disease Lifecycle of FHB	4
1.3 Disease Management Strategies for FHB	5
1.4 Causal Agents of FHB	7
1.5 Dynamics in the Population Structures of FHB Pathogens	10
1.6 The Genetics of FHB Pathogens and Toxin Production	12
1.7 Comparative Genomics as a Tool for Understanding FHB Pathogens	14
1.8 Thesis Outline	18

Chapter 2: Genomics of *Fusarium graminearum*, *F. asiaticum*, and *F. meridionale*

2.1 Abstract	21
2.2 Introduction	22
2.3 Results	24

2.3.1 Genome assembly information and statistics	24
2.3.2 The ten <i>Fusarium</i> pathogens that cause FHB are closely related	25
2.3.3 Regions of genomic variability in FHB fungi	27
2.3.4 Partitioning the pan genome into the common and accessory genomes for <i>Fusarium</i> species and isolates	29
2.3.5 Genetic variability in the common genes	30
2.3.6 Accessory genome reveals species and isolate-specific genes potentially involved in secondary metabolism and disease	31
2.4 Discussion and Conclusions	35
2.5 Methods	41
2.5.1 <i>Fusarium</i> isolates, sequencing, and genome assembly	41
2.5.2 Sequence analyses, comparisons, and annotation	42

Chapter 3: Intraspecies Interaction between Two *F. graminearum* Isolates Leads to Reduced Toxin Production and Disease

3.1 Abstract	56
3.2 Introduction	56
3.3 Results	58
3.3.1 Interference between <i>F. graminearum</i> isolates affects trichothecene production <i>in vitro</i>	58
3.3.2 Interference between <i>F. graminearum</i> isolates affects disease outcome <i>in planta</i>	60

3.3.3 Genome comparisons reveal differences in gene content between a 3-ADON and a 15-ADON producing isolate	63
3.3.4 Transcriptome profiling of axenic FG1 and FG2 isolates in mycotoxin-inducing conditions	64
3.3.5 Transcriptome profiling of co-inoculated cultures in mycotoxin-inducing conditions	65
3.4 Discussion and Conclusions	66
3.5 Methods	69
3.5.1 Fungal strains and inoculum preparation	69
3.5.2 Trichothecene production <i>in vitro</i>	70
3.5.3 Growth assays	70
3.5.4 Pathology testing in wheat and PCR/qRT-PCR analyses of isolates at the endpoint of infection	71
3.5.5 Genome sequencing, assembly, alignment, and gene prediction	72
3.5.6 RNA isolation and RNA sequencing	72
3.5.7 Normalization, validation, and analysis of RNA sequencing data ...	73

Chapter 4: Distribution of a Variable Region from the *Fusarium graminearum*

Genome

4.1 Abstract	86
4.2 Introduction	87
4.3 Results	88
4.3.1 Validation of the structure of contig_207 by PCR	88

4.3.2 Distribution and evolutionary history of genes from contig_207	89
4.3.3 Functional analysis of genes from contig_207	92
4.4 Discussion and Conclusions	94
4.5 Methods	97
4.5.1 Analyses of contig structure and distribution	97
4.5.2 Functional analyses of genes from contig_207	98

Chapter 5: Genomics of *Fusarium avenaceum*

5.1 Abstract	106
5.2 Introduction	106
5.3 Results	108
5.3.1 Genome structure and alignments of <i>F. avenaceum</i> genomes	108
5.3.2 Comparisons in gene content between <i>F. avenaceum</i> genomes	110
5.4 Discussion and Conclusions	111
5.5 Methods	113
5.5.1 Genome sequencing and comparative genomics	113

Chapter 6: Final Discussion

6.1 Summary of Findings	119
6.2 Studying FHB Pathogens in a Modern Genomics Era	123
6.3 Limitations of Largescale Genomics Studies	124
6.4 Closing Remarks	126

References	128
-------------------------	-----

LIST OF TABLES

Table 2.1 Description and Origin of <i>Fusarium</i> Isolates Used in this Study	51
Table 2.2 Genome and Assembly Statistics of <i>Fusarium</i> Isolates Used in this Study	52
Table 2.3 Isolate-Specific Genes Identified from Accessory Genome	53
Table 2.4 Secondary Metabolic Genes are Part of Accessory Genome	54
Table 2.5 Accessory Genes Contribute to Virulence, Toxin Production, and Development	55
Table 3.1 Genes Determined to be Differentially Expressed by Two-fold or Greater, p-value < 0.05	84
Table 3.2 Trichothecene Associated Genes have Increased Expression in FG2 Compared to FG1 by RNA-Seq	85
Table 4.1 Predicted Biological Roles for Genes from Contig_207	105
Table 5.1 Genome Statistics for <i>F. avenaceum</i> Genomes	118

Accompanying Materials: Tables S2.1 to 2.6 and Tables S3.1 to 3.4 are electronic files not included in the main text of the thesis.

LIST OF FIGURES

Figure 1.1 Fusarium Head Blight disease of wheat	19
Figure 1.2 Structures of trichothecene mycotoxin derivatives from FHB fungi in Canada	20
Figure 2.1 Concordance between genome alignments and taxonomy	44
Figure 2.2 Phylogeny of FHB pathogens by whole genome alignment	45
Figure 2.3 Sequence and gene content variability between genomes	46
Figure 2.4 The accessory genome is a subset of the pan genome	47
Figure 2.5 Sequence variability within common genes	48
Figure 2.6 Percent gene overlap between genomes	49
Figure 2.7 Top BLASTx hit of genes from the pan gene list not from PH-1or CS3005 . . .	50
Figure 3.1 HPLC chromatographs that show abundance of 3-ADON and 15-ADON in FG1 and FG2 in axenic and co-inoculated cultures	75
Figure 3.2 Co-inoculation of <i>F. graminearum</i> isolates reduces trichothecene production <i>in vitro</i>	76
Figure 3.3 Growth of FG1 and FG2 isolates in both solid and liquid media	77
Figure 3.4 Co-inoculation of <i>F. graminearum</i> isolates affects disease outcomes in wheat .	78
Figure 3.5 The FG1 isolate is more prevalent than the FG2 isolate at 21 dpi	80
Figure 3.6 Venn diagram of differentially expressed genes in co-inoculated cultures	81

Figure 3.7 Genes were determined to be differentially expressed in response to isolate interaction by both RNA-Seq and qRT-PCR	82
Figure 3.8 Trichothecene genes show a trend towards down-regulation in response to isolate interaction	83
Figure 4.1 Confirmation of the structure of contig_207 from FG1	99
Figure 4.2 Distribution of genes from contig_207 across the genomes of nine other closely related genomes from other FHB pathogens	100
Figure 4.3 A reference based mapping of sequencing reads from FG2 directly to contig_207 of FG1.	101
Figure 4.4 Relative copy number of the protease gene <i>FGSG_03432</i> or <i>g3557</i>	102
Figure 4.5 Percent overlap of genes from contig_207 in other <i>Fusarium</i> genomes	103
Figure 4.6 Normalized expression of genes from contig_207 after 96 h of infection in wheat	104
Figure 5.1 Genome alignment of the three <i>F. avenaceum</i> isolates (FAV1, FAV2, and FAV3) reveals that the genomes are highly syntenic	114
Figure 5.2 Supercontigs of the three <i>F. avenaceum</i> isolates; two from Canada (FAV1, FAV2) and one from Finland (FAV3)	115
Figure 5.3 Alignment of higher order assembly from FAV1 to <i>F. verticillioides</i> and <i>F. graminearum</i> chromosomes	116

Figure 5.4 Overlap in predicted genes for the three <i>F. avenaceum</i> genomes (FAV1, FAV2, and FAV3)	117
---	-----

LIST OF ABBREVIATIONS

3-ADON: 3-acetyldeoxynivalenol

15-ADON: 15-acetyldeoxynivalenol

AAFC: Agriculture and Agri-Food Canada

BAH: bromo-adjacent homology

bp: base pairs

cDNA: complementary DNA

Contig: Contiguous Sequence

DON: Deoxynivalenol

DNA: Deoxyribonucleic Acid

DPI: Days Post Inoculation

EF1 α : Elongation Factor 1Alpha

ELISA: Enzyme-Linked Immunosorbent Assay

FA: *Fusarium asiaticum*

FAV: *Fusarium avenaceum*

FG: *Fusarium graminearum*

FM: *Fusarium meridionale*

FHB: Fusarium Head Blight

GAPDH: Glyceraldehyde 3-Phosphate Dehydrogenase

HPLC: High Performance Liquid Chromatography

Indel: Insertion and/or Deletion

kb: kilo base pairs

KEGG: Kyoto Encyclopedia of Genes and Genomes database

Mb: mega base pairs

MIPS: Munich Information Center for Protein Sequences

NCBI: National Center for Biotechnology Information

NIV: Nivalenol

NMR: Nuclear Magnetic Resonance

NYN: Nedd4-BP1 YacP nucleases

ORDC: Ottawa Research and Development Center

OST-HTH/LOTUS: Oskar and vertebrate TDRD5/TDRD7

PCR: Polymerase Chain Reaction

QTL: Quantitative Trait Locus

RNA: Ribonucleic Acid

RQ: Relative Quantification

qRT-PCR: Quantitative Real-Time Polymerase Chain Reaction

RNA-Seq: Ribonucleic Acid Sequencing

SNP: Single Nucleotide Polymorphism

Tri: Trichothecene

Chapter 1: Introduction

1.1 Fusarium Head Blight (FHB): Impacts to the Economy and Human Health

Several fungal species from the genus *Fusarium* cause diseases in Canadian agricultural crops. A disease of particular interest is Fusarium Head Blight (FHB), which affects *Triticum* spp., including wheat. The disease appears as premature discolouration or bleaching of the wheat head or floral tissues (Trail 2009). In addition to physically damaging the crop and causing reduced yields as part of FHB infection, *Fusarium* fungi can produce a multitude of secondary metabolites that contaminate the wheat grains and make them unsafe for consumption by animals (Desjardins 2006; Hansen *et al.* 2015; Sieber *et al.* 2014). Common secondary metabolites include polyketides, non-ribosomal peptides, and terpenes (Hansen *et al.* 2015; Sieber *et al.* 2014). The terpene metabolite deoxynivalenol (DON) is of particular interest because it is highly toxic and its allowance in food products is regulated (Gilbert and Haber 2013). The economic impact of reduced yield and contamination with fungal toxins are substantial (Nganje *et al.* 2004). As such, there have been significant efforts from the scientific research community to better understand the factors contributing to host resistance, as well as pathogen disease and toxin production.

Contamination of crops by *Fusarium* pathogens is not a recent phenomenon; many FHB pathogens can infect a variety of crops and are likely to have been a common reoccurring pathogen of food crops, including wheat (Clear and Patrick 2010; Desjardins 2006; McMullen *et al.* 1997). Consumption of FHB infected wheat crops causes nausea and vomiting, and historical recordings of these disease symptoms may correlate with FHB prevalence even before the mechanism of disease was known (Desjardins 2006). In the early to middle 1900's, scientists

began a more thorough and systematic investigations of the fungi causing FHB and their associated toxins that are linked to disease symptoms. The non-contagious nature of the disease led scientists to discover that it was contamination of the crops with a food pathogen that was causing the disease (Desjardins 2006). In the 1980's, Canadian grains contaminated by fungi were screened for fungal contaminants including DON, and increased fungal infection by *Fusarium* correlated with increased toxin load (Desjardins 2006). Later, it was discovered that the disease symptoms associated with consumption of contaminated grains were specifically associated with the toxins that the fungi produce, specifically the trichothecene category of toxins, of which DON is a member (Desjardins 2006). Subsequent research has determined that DON is likely a protein synthesis inhibitor, capable of binding to the eukaryotic ribosomal subunit (Harris and Gleddie 2001).

By the end of the 20th century, the understanding of fungal toxins in contaminated food products and their associated health effects became better understood and policy makers began setting allowances for various fungal contaminants (Egmond and Jonkers 2004). In 2001, the World Health Organization first evaluated DON, making recommendations for its maximum intake, and later re-evaluated DON and its acetylated derivatives in 2011; they have set a provisional maximum tolerable daily intake of 1 µg/kg of body weight (World Health Organization and Food and Agriculture Organization of the United Nations 2011). Canada set an allowance for DON in food products of 2 mg/kg for unprocessed wheat and 1 mg/kg for baby food (Gilbert and Haber 2013).

Regular screening for DON has become standard practice in Canada, and to meet the standards for DON allowance, the Canadian agricultural industry blends together wheat from

different regions (Desjardins 2006; Gilbert and Tekauz 2000; Gilbert and Haber 2013). The prevalence of *Fusarium* species and toxins have been surveyed by the Canadian Grain Commission and the surveys indicate that between the years 1998 and 2005, the percentage of wheat samples contaminated by *Fusarium* ranged from 0.1 – 5.0% in Alberta, 2.6 – 29.0% in Saskatchewan, and 30.1 – 74.0% in Manitoba (Clear and Patrick 2010; Graefenham *et al.* 2014). This is in agreement with surveys of soil samples from 1994-1998 and sampling of grains from 2011-2013 that reported the greatest levels of disease and DON in samples that originated from Manitoba and Saskatchewan, and in some cases the DON levels exceeded the 2 mg/kg allowance (Clear and Patrick 2010; Graefenham *et al.* 2014). Although contaminated crops are blended to bring toxin levels into acceptable limits, the overall quality or grade of the crop decreases, thereby reducing its value on the market.

Canada is one of the top ten producers of wheat in the world, and since most of the wheat and its products are exported, overall yield and quality greatly impact the Canadian economy (Government of Canada 2010; Windels 2000). The allowance for *Fusarium* toxins is set by each nation and contamination of Canadian wheat can affect international trade if allowances are exceeded. Studies indicate that there are severe economic repercussions caused by FHB infection (Nganje *et al.* 2004). The destructive effects of FHB vary annually with corresponding effects on the economy (Clear and Patrick 2010; Graefenham *et al.* 2014; McMullen *et al.* 1997; Nganje *et al.* 2004; Windels 2000). In Canada, it is estimated that the losses by FHB infection were in the hundreds of millions of dollars in Quebec and Ontario (\$220 million 1990-2000), and in Manitoba (\$300 million from 1993-1998) (Windels 2000). A study that examined FHB infection in the United States of America from 1998-2000 determined disease-related losses to be in the hundreds of millions of dollars (\$160-457 million) annually

when considering production losses alone (Nganje *et al.* 2004). When indirect or secondary impacts of FHB are taken into account, losses were estimated to be as high as 1.4 billion dollars in a single year in the United States of America (Nganje *et al.* 2004).

1.2 The Disease Lifecycle of FHB

The lifecycle of FHB pathogens and stages of infection in wheat are well understood (Boenisch and Schäfer 2011; Trail 2009; Walter *et al.* 2010). An example of the wheat inflorescence tissues infected by FHB pathogens is shown in Figure 1.1. FHB infection begins when fungal spores land on mature wheat plants; inflorescence tissues are particularly susceptible (Trail 2009). These spores then germinate and the fungal hyphae begin to differentiate and form complex structures, such as compound appressoria that are used for host penetration (Boenisch and Schäfer 2011). These structures, along with degradative enzymes that the fungus secretes, degrade the structural barriers of the host and facilitate the fungus' entry into the host (Walter *et al.* 2010). Once inside the host, the fungus colonizes the apoplastic space and migrates towards the rachis and plant vascular tissues; DON has been reported to contribute to the spread of the pathogen and may play a role in fungal spread past the rachis node (Figure 1.1) (Boenisch and Schäfer 2011; Brown *et al.* 2010; Jansen *et al.* 2005). During initial infection and subsequent systemic colonization, degradative enzymes continue to play a role in the catabolism of structural components of the host and their conversion into nutrients for the fungus (Brown *et al.* 2010; Walter *et al.* 2010). As infection progresses, the fungus becomes more established in the host and begins to produce asexual spores, called macroconidia, on the exterior of the plant that disperses onto neighboring plant tissues (Trail 2009). The infected tissues become necrotic and eventually necrosis spreads throughout the host and the fungus alters its lifestyle to

proliferate on the tissue debris (Boenisch and Schäfer 2011; Brown *et al.* 2010). Until this point, the fungus had maintained its asexual lifestyle, producing conidiophores, but in the necrotized tissues, the fungus transitions into its sexual lifestyle (Trail 2009). Specialized sexual structures called perithecia are produced, which house sexual spores called ascospores. The ascospores will remain protected in the perithecia over the winter and are released in the spring, usually in conjunction with changes in weather, specifically light and moisture (Trail 2009). These spores then land on a susceptible host and the life cycle is perpetuated.

During initial infection, *F. graminearum* exhibits a biotrophic phase of fungal growth and does not kill the host directly, but as the fungus ramifies through the plant cells, the fungus begins to exhibit qualities of a necrotroph and forages on the dying plant materials. Pathogens such as *F. graminearum* that exhibit dual infection phases with an initial biotrophic lifestyle, followed by a necrotrophic lifestyle are called hemi-biotrophs (Ding *et al.* 2011).

1.3 Disease Management Strategies for FHB

The disease triangle is a classic model that describes FHB as an interaction between three variables: 1) the pathogen or disease causing agent, 2) the physical environment in which the disease occurs, and 3) the host. All three variables are required for disease and can be manipulated to some degree to manage or limit disease.

The most direct way to manage FHB is to target the pathogens that cause the disease by limiting the inoculum of spores or by targeting the fungus directly. FHB pathogens are reoccurring pests because they are globally distributed and are able to discharge their spores across great distances (Gilbert and Haber 2013, van der Lee *et al.* 2015). Strategies exist to

reduce the amount of spores available for causing infection; prominent cultural practices include tilling and crop rotation (Gilbert and Haber 2013; McMullen *et al.* 2012). Tilling or mechanically disrupting the soil can bury and increase the breakdown of plant debris, thereby reducing plant substrate for the fungus to overwinter and produce ascospores (McMullen *et al.* 2012). Crop rotation involves planting a non-host species between growing seasons, such as canola. Since FHB fungi can also infect barley and corn, it is recommended that these crops are not grown in rotation with wheat in regions that experience FHB infection (Gilbert and Haber 2013). Since the initiation of FHB coincides with ascospore discharge, some farmers plant wheat varieties with different flowering times to avoid the risk of having entire crops exposed to ascospores during flowering (McMullen *et al.* 2012). If the spores cannot be reduced or eliminated, a common and effective way of removing the pathogen from the disease triangle system is to use fungicides. Two types of fungicides are used with limited success, azoles, which inhibit the biosynthesis of ergosterol, a component of the fungal membrane, and strobilurin, which affects electron transport and energy pathways of the mitochondria (Edwards *et al.* 2001; Gilbert and Haber 2013). Biopesticides, biocontrols, or treatment of the soil with antagonistic microbes or chemicals may also reduce FHB prevalence (Gilbert and Haber 2013; McMullen *et al.* 2012). Forecasting tools are used to predict the inoculum load and severity of FHB in a given year, and the optimal time for farmers to apply pesticides (Gilbert and Haber 2013; McMullen *et al.* 2012).

The environment is also an important component of disease, but is much more difficult to control. It is known that FHB severity is greater in areas of increased rainfall (Clear and Patrick 2010; Graefenham *et al.* 2014). Temperature can also affect FHB, and can alter the species and population structure of FHB pathogens (Gilbert *et al.* 2014; van der Lee *et al.* 2015;

Zhang *et al.* 2010). Ideally, wheat should be grown in an environment not suitable or less suitable for FHB pathogen, such as areas of reduced rainfall.

Another effective way to limit the impact of FHB is to grow wheat cultivars that display resistance to FHB. Breeders have developed wheat varieties with different levels of resistance to FHB; unfortunately, many of these varieties do not always have the high yields or commercial traits that growers desire (Gilbert and Haber 2013, McMullen *et al.* 2012). FHB resistance is categorized into five different types: resistance to initial fungal infection (Type 1), resistance to fungal spread (Type 2), resistance to infection of the grain (Type 3), resistance to reductions in yield (Type 4), and resistance to trichothecenes or ability to metabolize trichothecenes (Type 5) (Rudd *et al.* 2001). Natural sources of resistances with quantitative trait loci (QTL) have been identified, but the genetic basis for resistance deployed through QTLs is not well understood (Buerstmayr *et al.* 2009). Natural resistance may be a multifactorial trait involving multiple host genes. These genes may also require certain environmental conditions or pathogens interactions to become active. This complicates the search for stable resistance to FHB.

1.4 Causal Agents of FHB

FHB is associated with a group of closely related fungi within the genus *Fusarium*. Many of the fungi that cause FHB can cause disease in other crops, particularly corn, barley, and rice (van der Lee *et al.* 2015). Fungi that cause FHB also produce different variants of the trichothecene mycotoxin and infect with different degrees of severity over their host range; some are even entirely non-pathogenic on some crops (O'Donnell *et al.* 2000; Starkey *et al.* 2007). The diverse nature of these pathogens has encouraged taxonomists to classify the fungi into different species and chemotypes. Early taxonomy of FHB pathogens organized species into several

lineages based on genealogy, which largely reflected species geography (O'Donnell *et al.* 2000; O'Donnell *et al.* 2004). In this early model, *Fusarium graminearum*, the dominant FHB pathogen, was structured as a species complex containing separate lineages or species, with the lineage from North America as lineage 7 (O'Donnell *et al.* 2000; O'Donnell *et al.* 2004). Lineage 1 was prominent in South America, while lineage 3 was found in Africa, and lineage 6 in Asia (O'Donnell *et al.* 2000). Many of these lineages are very similar morphologically, and can only be separated by molecular analysis with marker genes (Miedaner *et al.* 2008; O'Donnell *et al.* 2000; O'Donnell *et al.* 2004).

In Canada, surveys indicated that the causal agent of FHB differs between provinces and within growing regions of each province, with *F. graminearum* being the most dominant species (Clear and Patrick 2010; Graefenham *et al.* 2014; Gilbert and Tekauz 2000; Kelly *et al.* 2015). Other FHB pathogens, taxonomically related to *F. graminearum* include *F. culmorum*, *F. poae*, *F. sporotrichioides*, *F. cerealis* or *F. crookwellense*, and *F. pseudograminearum*. Another fungus *F. avenaceum* can also cause FHB, but is part of a more distant taxonomic group, the *F. tricinctum* species complex (Clear and Patrick 2000; Graefenham *et al.* 2014; Gilbert and Haber 2013; Kelly *et al.* 2015). Prior to the 1990's, surveys indicated that FHB was present in Western Canada, particularly in Manitoba (Clear and Patrick 2010). Manitoba and eastern regions of Saskatchewan are particularly suitable for FHB, likely due to increased moisture (Clear and Patrick 2000). In the 1990's and 2000's, disease outbreaks increased and FHB spread to other areas of Western Canada (Clear and Patrick 2000). Currently, in Manitoba and Ontario, the disease is dominated by *F. graminearum* (Clear and Patrick 2010; Graefenham *et al.* 2014). In Saskatchewan and Alberta, *F. graminearum* is also the dominant FHB species, but other species such as *F. culmorum* and *F. avenaceum* can be more prevalent depending on the growing region

and year (Clear and Patrick 2010; Graefenham *et al.* 2014). The Atlantic provinces also experience FHB, but detailed survey information is lacking compared to that of Western Canada, likely due to the fact that the Atlantic provinces produce < 1% of Canada's wheat (Government of Canada 2010). Different species of FHB pathogens can also coexist in the same geographical region, and therefore provide a diverse inoculum to compete for host resources during each growing season (Clear and Patrick 2010; Graefenham *et al.* 2014).

FHB is an agricultural problem that exists beyond Canadian borders and persists across the globe, with different species dominating each region. *F. graminearum* is the most common species reported and is the dominant species in North America (van der Lee *et al.* 2015). Both *F. graminearum* and *F. culmorum* are widely reported in Europe, with *F. graminearum* increasing its prevalence in recent years (Miedaner *et al.* 2008; van der Lee *et al.* 2015). In Australia, infection by *F. graminearum* occurs, but *F. pseudograminearum* is more commonly reported; although, in addition to FHB, crown rot disease of wheat by *F. pseudograminearum* is also of concern (Miedaner *et al.* 2008). In China, the dominant FHB pathogens are *F. graminearum* and *F. asiaticum*, and the disease has significant impacts on the world's largest producer of wheat (Zhang *et al.* 2010). In Asian regions, *F. graminearum* has host preference towards wheat, while *F. asiaticum* is more common on rice (van der Lee *et al.* 2015). In South America, reports indicate that *F. graminearum* and *F. asiaticum* are also common and show the same host preferences to wheat and rice, respectively, while *F. meridionale* is also prevalent and has host preference towards corn (van der Lee *et al.* 2015). Survey information in Africa is limited; however, *F. graminearum* is prevalent along with several other species, some of which have been discovered only recently (O'Donnell *et al.* 2008; van der Lee *et al.* 2015).

1.5 Dynamics in the Population Structures of FHB Pathogens

F. graminearum is the most globally distributed FHB pathogen and is the best understood. It is also the most reported FHB pathogen in Canada (Gilbert and Haber 2013). Interestingly, the emergence of *F. graminearum* as the dominant FHB pathogen in Canada may be a recent phenomenon. Surveys conducted prior to 1990 indicate that *F. avenaceum* was prevalent and has since been displaced by *F. graminearum* (Clear and Patrick 2000; Clear and Patrick 2010; Graefenham *et al.* 2014). Unlike most FHB pathogens in Canada, which are part of the *F. graminearum* species complex, *F. avenaceum* is part of the *F. tricinctum* species complex. *F. avenaceum* is genetically different from *F. graminearum*; it can cause disease on a greater variety of plant species, and has not been reported to make trichothecenes (Lysøe *et al.* 2014).

In Canada, *F. graminearum* has been separated into two distinct genetic populations; the distributions of these populations are dynamic, but one population is increasing in prevalence and distribution (Ward *et al.* 2008; Kelly *et al.* 2015). This emerging aggressive population also produces the trichothecene precursor 3-acetyldeoxynivalenol (3-ADON), while members from the less aggressive population produce 15-acetyldeoxynivalenol (15-ADON) (Figure 1.2) (Ward *et al.* 2008; Kelly *et al.* 2015). Although both of these precursors become deacetylated in the host to become deoxynivalenol, isolates can only produce one of the acetylated precursors, which is attributed to allelic differences in the gene *Tri8* (Alexander *et al.* 2011). The mechanisms by which 3-ADON producers have become more aggressive pathogens in wheat are unknown; however, environmental factors such as temperature and moisture may play a role (Gilbert *et al.* 2014). Other trichothecene types and precursors also exist in nature; for example, allelic variation in the gene *Tri1* is associated with the production of the alternate trichothecene NX-2 (Figure

1.2), while the genes *Tri13* and *Tri7* are associated with production of nivalenol and are non-functional in DON producers (Lee *et al.* 2002; Varga *et al.* 2015). Recent evidence indicates that NX-2 producers likely exist in Canada and may be within the genetic population that produces 3-ADON; thus far, no nivalenol producers have been reported in Canada (Kelly *et al.* 2015; Liang *et al.* 2015).

Population changes in FHB pathogens have also been observed outside of Canada, suggesting that dynamic shifts in population structures within species may be a global phenomenon. A similar shift between species towards *F. graminearum* as the dominant FHB pathogen has also occurred in Europe, which previously reported greater incidences of *F. culmorum* (van der Lee *et al.* 2015). Dynamics have also been observed within species in Asia where the dominant species is *F. asiaticum*. The species has also been reported to make different trichothecenes; namely, nivalenol (NIV), 3-ADON, or 15-ADON (Wang *et al.* 2011a; Zhang *et al.* 2010). In China, *F. asiaticum* typically produces 3-ADON or nivalenol (NIV) and 3-ADON producing strains may be displacing NIV producing strains in some regions (Zhang *et al.* 2010).

Surveys have determined that many globally distributed species that cause FHB may be spreading to new regions. For instance, *F. asiaticum* was recently detected in North America and similarly, *F. graminearum* has seen an increase in Asia (Miedaner *et al.* 2008; Zhang *et al.* 2010). FHB pathogens are known for their ability to disperse their spores across great distances; however, globalization and trade also presents opportunities for pathogens to spread to new areas (Gilbert and Haber 2013; van der Lee *et al.* 2015). It is unclear if the recent migration dynamics of FHB species and populations are the result of increasing trade between countries or due to

change in environments and agricultural practices (Gilbert and Haber 2013; Kelly *et al.* 2015; O'Donnell *et al.* 2008; van der Lee *et al.* 2015).

1.6 The Genetics of FHB Pathogens and Toxin Production

The economic and health impacts of FHB fungi prompted the genome sequencing of *F. graminearum* isolate PH-1, which provided an important foundation for the genetic dissection of the fungal mechanisms surrounding disease and toxin production (Cuomo *et al.* 2007). The genome is approximately 36 Mb in length, and contains very few repetitive elements; it has been suggested that the species has an effective repeat-induced point mutation mechanism that eliminates repetitive elements (Cuomo *et al.* 2007). When compared to other ascomycete fungi, the genome has more transcription factors, hydrolytic enzymes, and transmembrane transporters (Cuomo *et al.* 2007). The genome sequence is also under constant revision by the scientific community and was recently re-sequenced to better resolve regions in the telomeres and centromeres, which has resulted in alternate genome assemblies (King *et al.* 2015). Updates to gene entries are ongoing and are cross-referenced with data from predictive software for protein localization and gene function; however, more than half of genes, ~52.3%, are functionally unclassified (Wong *et al.* 2011). Nevertheless, many of the biochemical pathways within the fungus have been mapped by the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (<http://www.kegg.jp/kegg/pathway.html>). The genome has also facilitated other largescale studies, such as the computational derivation of protein-protein interaction networks and metabolic gene clusters (Sieber *et al.* 2014; Zhao *et al.* 2009). The genome has also allowed for largescale gene expression analyses by microarray and RNA sequencing, which have been useful to identify gene expression patterns associated with developmental and physiological

phenomenon in the fungus, including genes activated during the infection in different hosts (Dash *et al.* 2012; Lysøe *et al.* 2011; Subramaniam *et al.* 2015; Zhao *et al.* 2013). Gene expression datasets have also enabled the construction of a gene regulatory network that incorporates both transcriptional regulators and signalling proteins (Guo *et al.* 2016; Lawler *et al.* 2013).

Largescale genetic studies have also allowed us to identify pathways involved in disease and toxin production. Several studies have globally targeted entire gene categories, including nearly all kinases, phosphatases, and transcription factors; several from each gene category have been implicated in toxin production and disease in wheat (Son *et al.* 2011, Wang *et al.* 2011b; Yun *et al.* 2015). More directed analyses have identified genes that both positively and negatively affect DON production and disease in wheat (Gardiner *et al.* 2009; Hou *et al.* 2002; Subramaniam *et al.* 2015; Walkowiak and Subramaniam 2014; Wang *et al.* 2014). Since the availability of the first draft genome, nearly one-tenth of genes have been studied by reverse genetics and it is possible in the near future, that a systems based approach can be undertaken to integrate pathways associated with toxin production and disease to various environmental conditions and the genetic background of the fungus (Jia and Tang 2015).

To date, the only virulence factor that has been well described to contribute to FHB infection has been the trichothecene mycotoxins; as such, there has been significant research into understanding the genetics and biosynthetic pathways involved in trichothecene production (Jansen *et al.* 2005). Most of the genes involved in trichothecene production have been identified and twelve are localized in a gene cluster, while four genes are located outside of the cluster (Kimura *et al.* 2007). Within the main cluster are the two genes that encode positive

transcriptional regulators, *Tri6* and *Tri10* (Nasmith *et al.* 2011). More detailed analyses of *Tri6* have discovered that it may regulate pathways outside of DON biosynthesis, including primary metabolic pathways (Nasmith *et al.* 2011, Subramaniam *et al.* 2015). In addition to containing regulatory genes, the cluster also contains a transporter gene, *Tri12* (Kimura *et al.* 2007). The other genes within the gene cluster encode biosynthetic enzymes, including the terpene synthase gene *Tri5* that catalyzes the first committed step in the pathway (Kimura *et al.* 2007). A subset of the biosynthetic genes within the trichothecene gene cluster exhibit genetic variation, including *Tri8*, *Tri7*, and *Tri13*. These genes have been demonstrated to be non-functional or have allelic differences in different populations of FHB pathogens, and these differences contribute to structural differences in the toxin molecule (Alexander *et al.* 2011; Lee *et al.* 2002; Varga *et al.* 2015). For instance, alternate alleles of *Tri8* result in differences in the deacetylation of 3,15-diacetylDON, resulting in either 3-ADON or 15-ADON (Alexander *et al.* 2011). Functional copies of *Tri13* and *Tri7* are required for production of nivalenol and modifications thereof, but not for the synthesis of DON and its derivatives (Lee *et al.* 2002). Variation in *Tri1*, a gene outside of the gene cluster also contributes to structural variation in the trichothecene molecule and can result in the production of NX-2, which is missing a ketone group on the toxin backbone (Varga *et al.* 2015). Other genes involved in trichothecene production but located outside of the gene cluster include the biosynthetic genes *Tri16*, *Tri101* and *Tri201* (Kimura *et al.* 2007).

1.7 Comparative Genomics as a Tool for Understanding FHB Pathogens

Studies have determined that FHB pathogens are a diverse set of fungi with clear differences in pathogenicity, toxin production, host specificity, and growth, and that these phenotypic variations occur both within and between species and genetic populations (Bai 1996;

O'Donnell *et al.* 2000; Ward *et al.* 2008; van der Lee *et al.* 2015; Zhang *et al.* 2010). Some isolates have been reported as being non-pathogenic on wheat (O'Donnell *et al.* 2000). It is conceivable that many of these differences are associated with adaptation to specific niches or environments (Gilbert *et al.* 2014; Zhang *et al.* 2010; van der Lee *et al.* 2015). The genetic bases for some of the differences in toxin production are known; for example, differences in acetylation and hydroxylation of trichothecenes are associated with modifications or disruption of the genes *Tri8*, *Tri7*, and *Tri13* (Alexander *et al.* 2011; Lee *et al.* 2002). There is a growing concern in Canada surrounding the emergence of more aggressive populations of *F. graminearum* that produce the 3-ADON trichothecene variant, which may be displacing 15-ADON producing strains of the fungus that belong to a separate genetic population (Ward *et al.* 2008; Kelly *et al.* 2015). Evidence indicates that the distribution of these different chemotypes is changing, possibly due to evolution and adaptation to changes in the environment (Ward *et al.* 2008; Gilbert *et al.* 2014). Since pathogenicity variants are observed within populations of 3-ADON producing *F. graminearum* strains, it is postulated that there are other factors, unrelated to trichothecenes that are involved in the emergence of more aggressive variants (O'Donnell *et al.* 2000; Ward *et al.* 2008). The available genome of the *F. graminearum* isolate PH-1 has significantly added to our understanding of the genetics of FHB pathogens and has facilitated the identification of genes and pathways that contribute to disease and toxin production; nonetheless, additional genomes are required to fully understand the diverse nature of FHB pathogens. Comparative genomics has been extensively used to study the biological diversity of plant pathogenic fungi and would therefore be an appropriate approach to better understand the biology of FHB pathogens.

Comparative genomics approaches have been used to study species differences in fungi, but only recently, largescale studies have been done to study diversity within the same species. The first draft sequence of fungal genome, *Saccharomyces cerevisiae* was released in 1996 (Goffeau *et al.* 1996). Subsequently, the availability of multiple fungal genomes has allowed scientists to compare genomes between different fungal species. In 2008, a comparison of 36 genomes from filamentous fungi and oomycetes, including seven from plant pathogens, revealed a list of genes and gene families that are more likely to occur in plant pathogens (Soanes *et al.* 2008). A direct comparison between the biotrophic pathogen *Hyaloperonospora arabidposis* and two hemibiotrophic *Phytophthora sp.* revealed important genetic differences in the infection strategies between theses oomycetes (Baxter *et al.* 2010). Whole genome sequencing and gene expression analyses of four hemibiotrophs from the genus *Colletotrichum* revealed virulence genes associated with either biotrophic or necrotrophic phases of infection (Gan *et al.* 2013; O'Connell *et al.* 2012). Another study comparing the genomes of two *Verticillium sp.* and *F. oxysporum*, all of which cause vascular wilt disease, revealed genes and gene islands that are important for pathogenicity and niche adaptation to this disease (Klosterman *et al.* 2011). A direct comparison of the genomes of two maize pathogens, *Ustilago maydis* and *Sporisorium reilianum*, revealed that regions that are not conserved between genomes are enriched in gene clusters and other virulence related genes (Schirawski *et al.* 2010). Other studies also revealed features such as repetitive sequences, telomeres, and conserved syntenic blocks that have been shown to contain genes associated with pathogenicity (Xu *et al.* 2006).

As of April 2016, there are over 1500 fungal genome assemblies available in the National Center for Biotechnology Information (NCBI) database, over 50 of which correspond to *Fusarium spp.* With the use of next-generation sequencing technologies, the number of available

genomes has grown significantly, including genomes within the genus *Fusarium*, as well as genomes within the same species complex. Although many of these genomes have been released as drafts, some genomes have been investigated in great detail; for example, the genomes of *F. fujikuroi*, *F. pseudograminearum*, and *F. langsethiae* are well described; the latter two cause FHB (Gardiner *et al.* 2012; Lysøe *et al.* 2016; Wiemann *et al.* 2013).

The first available and best studied genomes for *Fusarium* species correspond to the *F. graminearum*, *F. verticillioides*, *Nectria haematococca* (*F. solani*), and *F. oxysporum*. A detailed comparison of the *F. graminearum*, *F. verticillioides* and *F. oxysporum* genomes revealed chromosome fusion and translocation events, local rearrangements, as well as many features unique to the species (Ma *et al.* 2010). Approximately, 40% of the *F. oxysporum* genome is unique and not part of the core *Fusarium* genome, most of which exists in lineage-specific chromosomes that are similar in some respects to the gene islands found in bacteria. These variable regions, outside of the core genome, are of particular interest for pathogenicity related genes. Transfer of lineage specific chromosomes, which converted a non-pathogenic *F. oxysporum* strain into a pathogenic strain, confirmed that lineage specific regions contain pathogenicity related genes in *Fusarium* and are important for infection and niche specialization (Ma *et al.* 2010). One of the transferred lineage specific chromosomes contained several virulence factors known to be involved in infection of tomato (Ma *et al.* 2010). Given the diversity of *F. oxysporum*, core genes and important species-specific regions are difficult to resolve without multiple genome sequences. Genome sequencing of 10 additional *F. oxysporum* isolates from infected cabbage, radish, banana, pea, melon, and cotton, as well as an isolate from human blood and a biocontrol agent are complete or currently being assembled by the Broad Institute (<https://www.broadinstitute.org/>) and should facilitate the characterization of important

genes involved in host specificity and virulence. Of particular importance is the *F. oxysporum* pathogen of banana, which causes the devastating Panama disease and is a serious threat to the future of banana cultivation (Guo *et al.* 2014). The genome sequences of several of these isolates are currently publically available in the Broad Institute and NCBI databases and have been used to identify genes associated with disease (Guo *et al.* 2014; Ma *et al.* 2014; Schmidt *et al.* 2016). Similarly, sequencing and comparative genomics of several *F. oxysporum* strains that infect melon identified genes that interact with the host defense response pathways leading to either susceptible or resistant disease outcomes (Schmidt *et al.* 2016). Similar approaches of comparative genomics could easily be applied to FHB pathogens.

1.8 Thesis Project Outline

The goal of this thesis is to characterize the genetic diversity of FHB pathogens in Canada and abroad. This will be accomplished using a combination of comparative genomics, metabolic profiling, and virulence/toxin analyses. In Chapter 2, we sequence and compare the genomes of multiple FHB pathogens and provide insights into the genetic similarities and differences between *F. graminearum* and its close relatives within the *F. graminearum* species complex, namely *F. meridionale* and *F. asiaticum*. Chapter 3 focuses more specifically on differences between two *F. graminearum* isolates, a 3-ADON and a 15-ADON producer, and the consequences of intraspecies interactions in culture and during host infection. Chapter 4 concentrates on the analysis of a single contiguous sequence from the genome of a 3-ADON producing *F. graminearum* strain, which is highly variable when compared to other *F. graminearum* genomes. Chapter 5 addresses the genomics of a more distantly related FHB pathogen, *F. avenaceum*, which is part of a different *Fusarium* species complex, *F. tricinctum*.

Figure 1.1: Fusarium Head Blight disease of wheat. The susceptible wheat cultivar ‘Roblin’ was inoculated with 1000 *F. graminearum* spores between the palea and the lemma; the inoculated spikelet is labelled with a red dot. The image was taken 10 days post inoculation (dpi).

Figure 1.2: Structures of trichothecene mycotoxin derivatives from FHB fungi in Canada. Structures of the trichothecene mycotoxin derivatives 15-acetyldeoxynivalenol, 3-acetyldeoxynivalenol and NX-2 are presented. Acetyl (blue box) and ketone (red box) groups are variable and their variability is attributed to allelic differences in the genes *Tri8* and *Tri1* respectively.

Chapter 2: Genomics of *Fusarium graminearum*, *F. asiaticum*, and *F. meridionale*

2.1 Abstract

The *Fusarium graminearum* species complex is composed of many distinct fungal species that cause several diseases in economically important crops, including Fusarium Head Blight of wheat. Despite being closely related, these species and individuals within species have distinct phenotypic differences in toxin production and pathogenicity, with some isolates reported as non-pathogenic on certain hosts. In this chapter, we compare genomes and gene content of eight newly sequenced isolates from the species complex, including the first available genomes of *F. asiaticum* and *F. meridionale*, with two other genomes reported in previous studies. A comparison of genome structure and gene content revealed a 93-99% overlap across all ten genomes. We identified more than 700kb of single nucleotide polymorphisms (SNPs), insertions, and deletions (indels) within common regions of the genome, which validated the species and genetic populations reported within species. We constructed a non-redundant pan gene list containing 15297 genes from the ten genomes; of these, 1827 genes or 12% were absent in at least one genome. Genes that were absent between genomes were co-localized in telomeric regions and select regions within chromosomes with a corresponding increase in SNPs and indels. Many of the absent genes between isolates are predicted to encode proteins involved in secondary metabolism and other functions associated with disease. The genomic resources we have contributed will be useful for the identification of genes that contribute to the phenotypic variation and niche specialization that have been reported between members of the *F. graminearum* species complex.

2.2 Introduction

Fungi belonging to the genus *Fusarium* are very diverse and occupy a wide variety of ecological niches. The genus is divided into several distinct species complexes, each with many individual species (Geiser *et al.* 2013; O'Donnell *et al.* 2000). In this study, we focused our attention on closely related individuals from the *Fusarium graminearum* species complex, which cause the Fusarium Head Blight (FHB) disease of wheat. This species complex is composed of several distinct fungal species, including but not limited to *F. graminearum*, *F. asiaticum*, and *F. meridionale* (Geiser *et al.* 2013). The genome of *F. graminearum* isolate PH-1 was the first sequenced from the species complex, and has been revisited as additional information on gene and genome structure has become available (Cuomo *et al.* 2007; King *et al.* 2015; Wong *et al.* 2011; Zhao *et al.* 2013). Large-scale genomics studies have used this genome as a reference to study different biological phenomena, such as genome-wide association analyses of SNPs, as well as analyses of recombination events by restriction site associated DNA sequencing (Talas and McDonald 2015; Talas *et al.* 2016). Few studies have explored the genetic diversity of members of the *F. graminearum* species complex by performing whole genome assembly and comparative genetics; these studies have also largely focused on *F. graminearum sensu stricto* – a single lineage within the complex (Cuomo *et al.* 2007; Gardiner *et al.* 2014).

As part of the infection process, these fungi produce a harmful trichothecene mycotoxin that contributes to disease progression and accumulates in infected plant tissues (Jansen *et al.* 2005). Although genetic differences in genes that contribute to trichothecene variations are known, our current knowledge of other factors that lead to variations in disease profiles remains largely unknown (Alexander *et al.* 2011; Lee *et al.* 2002; Varga *et al.* 2015). Information gathered from sequences of additional members of the *F. graminearum* species complex will

lead us to better understand the genetic variations between isolates that are associated with reported phenotypic diversity.

This study reports on genetic differences between ten closely related members of the *Fusarium graminearum* species complex that cause FHB. Eight isolates were sequenced as part of this study (Accessions: NRRL6101, NRRL28720, NRRL28721, NRRL28723, NRRL28336, DAOM180378, DAOM233423, and DAOM241165) and two were sequenced previously (PH-1 and CS3005) (Cuomo *et al.* 2007; Gardiner *et al.* 2014). NRRL6101 was isolated from barley in Japan, which was previously categorized as *F. asiaticum* and was reported to be non-pathogenic in wheat (O'Donnell *et al.* 2000; Vesonder *et al.* 1982). NRRL28720, NRRL28721, and NRRL28723 were isolated from maize in Nepal. While NRRL28720 belongs to the species *F. asiaticum*, NRRL28723 is of the related species *F. meridionale*. The species classification for NRRL28721 is still unclear and has been suggested to be a species hybrid (*F. asiaticum* x *F. meridionale*) (Carter *et al.* 2002; O'Donnell *et al.* 2000; O'Donnell *et al.* 2004; Starkey *et al.* 2007; Ward *et al.* 2002; Yang *et al.* 2008). Further investigation with additional markers suggested that this isolate may be in a basal group within the *F. asiaticum* clade or belong within *F. meridionale* (O'Donnell *et al.* 2004; Starkey *et al.* 2007). The isolates PH-1, DAOM233423, DAOM241165, DAOM180378, NRRL28336, and CS3005 belong to *F. graminearum* species. With the exception of CS3005, which was isolated in Australia, all other isolates were isolated in North America (Cuomo *et al.* 2007; Gardiner *et al.* 2014; O'Donnell *et al.* 2004; Walkowiak *et al.* 2015). PH-1 and DAOM180378 were isolated from maize, NRRL28336, DAOM233423, and DAOM241165 were isolated from wheat, and CS3005 was isolated from barley. Both the trichothecene chemotype and pathology on wheat for each isolate has been reported (Table 2.1).

We compared the genome sequences of these ten closely related *Fusarium* isolates and identified genes that are part of the common genome (present in all isolates), as well as the accessory genome, which are potentially involved in niche specialization within and between species.

2.3 Results

2.3.1 Genome assembly information and statistics

We performed whole genome sequencing of eight fungal isolates (NRRL6101, NRRL28336, NRRL28720, NRRL28721, NRRL28723, DAOM180378, DAOM233423, and DAOM241165) from the *F. graminearum* species complex that are associated with FHB disease in plants (Table 2.1). The newly sequenced genomes, along with two previously sequenced isolates, were selected based on reported differences in geographical distribution, taxonomy/species, host isolation, toxin production and pathology (O'Donnell *et al.* 2000; O'Donnell *et al.* 2004; Ward *et al.* 2002; Starkey *et al.* 2007; Yang *et al.* 2008) (Table 2.1). Overall, the ten sequenced genomes were comparable in size and structure (Table 2.2). The sizes of the ten genomes ranged from approximately 36.4 Mb to 36.7 Mb. The N50 values, which represent the length of the contig at which 50% of the genome is contained in a contig of that size or greater, ranged from 0.11 Mb to 1.21 Mb for the high-throughput sequenced genomes; the genome coverages ranged from 32x to 144x (Table 2.2).

2.3.2 The ten *Fusarium* pathogens that cause FHB are closely related

To determine the degree of overlap between the ten genomes, we performed genome alignments between the ten genomes from Table 2.2, as well as genomes from 41 other plant pathogenic fungi obtained from NCBI (National Center for Biotechnology Information) using two different algorithms that are part of the MUMMER software package, NUCMER and PROMER (Figure 2.1) (Kurtz *et al.* 2004).

A direct pairwise comparison with the stringent aligner NUCMER, which is designed for the alignment of closely related sequences, revealed at least 96.4% sequence alignment between the six *F. graminearum* genomes (PH-1, CS3005, DAOM233423, DAOM241165, DAOM180378, and NRRL28336). Similarly, when the two *F. asiaticum* genomes (NRRL6101 and NRRL28720) were compared, 97.7% of the genomes aligned. Lastly, a comparison between the two *F. meridionale* genomes (NRRL28721 and NRRL28723) revealed a 99.0% alignment between them (Figure 2.1). When we aligned the genomes of different species, *F. graminearum* and *F. asiaticum* had at least 93.1% sequence alignment, *F. graminearum* and *F. meridionale* had at least 94.4% sequence alignment, and *F. asiaticum* and *F. meridionale* had at least 93.8% sequence alignment. The ten genomes had between 90.9-92.0% alignment to *F. pseudograminearum* and 64.9-65.5% alignment to *F. langsethiae*, but only 3.5-12.8% aligned to other *Fusarium* genomes, and < 1% alignment to genomes from other plant pathogenic fungi (Figure 2.1).

Analysis with PROMER, which is designed for the alignment of distantly related sequences, yielded results similar to those of NUCMER in pairwise comparisons of the ten genomes to one another and to *F. pseudograminearum* and *F. langsethiae* (Figure 2.1). However,

the alignment to other *Fusarium* species increased from 3.5-12.8% to 39.0-50.8%. Alignment to other phytopathogen Ascomycetes increased from <1% to 3.4-26.1%, and alignment to phytopathogen Basidiomycetes remained lower at 0.8-3.9% (Figure 2.1). The concordance observed between genomes from our pairwise alignments supports taxonomic phylogenies that suggested that *F. graminearum*, *F. asiaticum*, and *F. meridionale* are closely related, while *F. pseudograminearum* and *F. langsethiae* are distantly related, and that other fusaria belong to more distant taxonomic groups, followed by other genera and other phyla (Geiser *et al.* 2013; O'Donnell *et al.* 2004; Starkey *et al.* 2007; Kristensen *et al.* 2005).

In addition to analyzing genetic overlap by direct pairwise alignments, we also identified single nucleotide polymorphisms (SNPs) and insertions/deletions (indels) in all genomes, which were used to construct a maximum likelihood phylogenetic tree (Figure 2.2). We used PANSEQ to identify a total of 704566 SNPs and indels that differed between the ten genomes (Laing *et al.* 2010). As observed, the ten isolates arranged according to their previously reported species assignments and were in complete agreement with the pairwise alignment data describing genetic overlap (Figure 2.1). Together, the data indicate that PH-1, DAOM233423, DAOM180378, DAOM241165, CS3005, and NRRL28336 are *F. graminearum*, while NRRL6101 and NRRL29720 are *F. asiaticum* (Geiser *et al.* 2013; O'Donnell *et al.* 2004; Starkey *et al.* 2007; Kristensen *et al.* 2005) (Figure 2.2). The phylogeny also supports our pairwise genome alignment data which indicated that NRRL28721 is closely related to NRRL28723 and likely corresponds to *F. meridionale* (Figure 2.1). Interestingly, the two *F. graminearum* isolates that produce the trichothecene precursor 3-ADON (DAOM241165 and NRRL28336) form a distinct genetic group from the four isolates that produce 15-ADON (PH-1, DAOM233423, DAOM180378, and CS3005) (Figure 2.2). This supports claims that individuals from these two

chemotypes (3-ADON and 15-ADON) belong to distinct separate genetic populations (Kelly *et al.* 2015; Liang *et al.* 2015). The *F. graminearum* isolate CS3005, originating from Australia arranged in the same phylogenetic group as the other 15-ADON producers from North America, implying that it may belong to the same genetic population (Figure 2.2).

2.3.3 Regions of genomic variability in FHB fungi

To visualize the genomic regions with sequence differences, the genomes were organized into 100kb sections and the average number of SNPs and indels to the genomes of each species were plotted across the length of the four chromosomes (Figure 2.3). Overall, we observed that the average number of SNPs and indels within species were fewer than between species. For example, when the PH-1 genome was used as a reference, the average number of SNPs and indels within the genomes of *F. graminearum* species (black line, Figure 2.3A) were fewer than SNPs and indels identified in either *F. meridionale* (blue line, Figure 2.3A) or *F. asiaticum* (red line, Figure 2.3A) genomes. Similar to the observation with *F. graminearum* species, when a genome from a separate species was used as a reference (Figure 2.3B and C), we observed that SNPs and indels were fewer in genomes within the same species than between genomes from different species. Together, we observed 105-189kb of SNPs and indels when comparing genomes within the same species and 796-893kb of SNPs and indels when comparing genomes between *F. graminearum*, *F. asiaticum*, and *F. meridionale*; however, the locations of nucleotide variations were similar in each case. Sequence variation was co-localized at the ends of chromosomes and select regions within chromosomes, similar to what has been reported previously (Figure 2.3) (Cuomo *et al.* 2007). In addition, we also observed islands of increased nucleotide variation in specific genomes. For example, in the genomes of CS3005, NRRL28336,

and NRRL28720, we observed regions where the nucleotide variations increased compared to most other genomes (boxed, Figure 2.3).

There is preponderance of evidence to suggest that niche adaptation, toxin production, and disease outcomes are attributed to differences in gene content and variation within genes (Alexander *et al.* 2011; Hansen *et al.* 2015; Lee *et al.* 2002; Ma *et al.* 2010; Proctor *et al.* 2009; Sieber *et al.* 2014; Walkowiak *et al.* 2015). We were interested to know if the nucleotide variations observed were associated with absent genes between species. To assess the genetic variations both within and between species, we compared the gene content between the ten genomes. Gene sequences for isolates PH-1 and CS3005 were available from previous studies (Gardiner *et al.* 2014; Walkowiak *et al.* 2015; Wong *et al.* 2011). AUGUSTUS was used to predict genes from the eight genomes sequenced in this study (Stanke and Waack 2003). Genes were considered absent if they had the expected value $>1E-10$ by BLASTn compared to other genomes. Our analysis showed that regardless of the reference genome, absent genes in *F. graminearum* (black symbols), *F. meridionale* (blue symbols), and *F. asiaticum* (red symbols) were all located in regions with high SNPs and indels (Figure 2.3). In addition to absent genes, we also identified genes that were unique to each isolate (purple symbols), which were also localized to regions of increased nucleotide variability (Figure 2.3).

2.3.4 Partitioning the pan genome into the common and accessory genomes for *Fusarium* species and isolates

The pan gene list was started by including all PH-1 genes, then genes from other genomes were compared to the current pan gene sequences with BLASTn and added if they had expected values (e-value > 1E-10). This non-redundant gene list contained a total of 15297 genes that are present in at least one of the ten *Fusarium* genomes, some of which were common to all genomes, unique to a species, or unique to an isolate (Accompanying Materials: Table S2.1). Of the 15297 genes in the pan genome, 12% or 1827 genes form the accessory genome (Figure 2.4, Accompanying Materials: Table S2.1). The accessory genome contributes to phenotypic differences and might encode proteins involved in biochemical pathways and functions that are not essential for *Fusarium* growth but confer selective advantages, such as adaptation to different niches or colonization of a new host (Laing *et al.* 2010).

The accessory genome was further parsed into genes that are unique to a species or each isolate (Accompanying Materials: Table S2.2). Analyses identified a total of 105 species-specific genes that were present in all six *F. graminearum* isolates, but were absent in all genomes of the other two species (Figure 2.4). Similarly, a total of 126 and 184 species-specific genes were identified in *F. meridionale* and *F. asiaticum*, respectively (Figure 2.4).

Analysis of the accessory genome also revealed genes that are present in a single isolate, and absent in all others (Table 2.3, Accompanying Materials: Table S2.2). We observed from 12-92 unique genes in each of the ten isolates (Uniquely Present, Table 2.3, Accompanying Materials: Table S2.2). In contrast, the *F. asiaticum* isolate NRRL6101 had the greatest number

of “uniquely absent” genes (34), absent in one genome and present in all others, and this isolate has been reported to be unable to infect wheat (O’Donnell *et al.* 2000).

2.3.5 Genetic variability in common genes

Since all ten of the fungi under investigation are closely related cereal pathogens, we wanted to identify common genes to all ten isolates that may be variable and therefore contribute to phenotypic differences. To identify those genes that are common between genomes, but are highly variable, we used SNP and indel data from the whole genome alignments. We identified 9269 common genes with ≤ 40 SNPs or indel sites per kilo base pairs (kb) and 163 genes that had >200 sites per kb (Figure 2.5, Accompanying Materials: Table S2.3). Since 110 of the 163 genes originated from PH-1, we used the MIPS FunCat database (<http://mips.helmholtz-muenchen.de/funcatDB/>), designed using the genetic data for PH-1, to detect functional enrichment of gene categories (Accompanying Materials: Table S2.4). The analysis revealed that 81 (73.6%) are functionally unclassified, compared to 52.3% for all genes from PH-1. Twenty nine genes that were functionally classified were enriched (p-value < 0.05) in a wide array of cellular processes including motility, energy, metabolism, and cell migration (Accompanying Materials: Table S2.4). This included *FGSG_03532* or *Tri8*; allelic variation in *Tri8* is attributed to the production of either 15-ADON or 3-ADON (Alexander *et al.* 2011). In addition, three putative transcription factors *FGSG_13457*, *FGSG_08954*, and *FGSG_10508* were also identified. These three genes have been previously targeted for functional analysis, and disruption of *FGSG_13457* resulted in no apparent phenotypic difference in toxin production or disease, while no mutants were obtained for *FGSG_08954* and *FGSG_10508* (Son *et al.* 2011). Two nodulin-like precursor proteins, *FGSG_03381* and *FGSG_03550* were also identified. Proteins with nodulin-like domain are often found in plants and are associated with plant-

microbe interaction (Denancé *et al.* 2014). Lastly, three heterokaryon incompatibility genes *FGSG_10601*, *FGSG_08120* and *FGSG_08144* with roles in non-self-recognition were also identified (Glass and Kaneko 2003; Paoletti *et al.* 2007).

2.3.6 Accessory genome reveals species and isolate-specific genes potentially involved in secondary metabolism and disease

Accessory genes were analysed for their role in niche adaptation, speciation, toxin production, and disease. For example, there were 105 *F. graminearum*-specific genes, 90 of which originated from PH-1 and functional enrichment analyses of these by the MIPS FunCat database revealed that 72 or 80% of these were unclassified; however, the 18 genes that were functionally classified were enriched in metabolic and defense pathways (Accompanying Materials: Table S2.5). The genes involved in defense included two aldehyde dehydrogenases, *FGSG_17538* and *FGSG_01759*, as well as gene *FGSG_17130* encoding for a protein similar to pisatin demethylase cytochrome P450. Pisatin is a phytoalexin whose detoxification has been demonstrated by the action of demethylase genes from *F. solani* (Milani *et al.* 2012). Similarly, there were 184 genes that were specific to *F. asiaticum* (Table 2.4). Among these included secondary metabolic genes such as a non-ribosomal peptide synthetase APS1, involved in apicidin biosynthesis and a polyketide synthase PKS40 associated with W493 biosynthesis (Table 2.4) (Hansen *et al.* 2015; Sørensen *et al.* 2014). There were also 126 genes specific to *F. meridionale*, including a gene *g3644* from NRRL28723 that was annotated as a ‘pectin lyase fold virulence factor’ implicated in cell wall catabolism (Accompanying Materials: Table S2.6) (Mayans *et al.* 1997). Twelve heterokaryon or vegetative incompatibility genes were also identified as being species-specific; these gene classes are known to diversify in fungi and may have roles in speciation (Glass and Kaneko 2003; Paoletti *et al.* 2007).

Next, our analysis was extended to isolate-specific genes. Of the 1827 genes from the accessory genome, between 12 and 92 genes were present in a single isolate and were considered unique to that isolate (Table 2.3). Unique genes included genes of both primary and secondary metabolism, as well as signal transduction. Example includes a glutamine synthase, *g9634* from NRRL28336, and a non-ribosomal peptide synthase, *g11745* from NRRL6101 (Table 2.4; Accompanying Materials: Table S2.6), involved in primary and secondary metabolism, respectively. Furthermore, among the 92 genes unique to NRRL28336, nine may possess kinase function and may be involved in signal transduction (Accompanying Materials: Table S2.6). A kinase unique to PH-1, *FGSG_11614*, was disrupted and resulted in reduced DON production (Wang *et al.* 2011b).

Genes that were uniquely absent in an isolate were also analysed (uniquely absent, Table 3). The isolate NRRL6101 reported to have limited virulence in wheat contained the greatest number of uniquely absent genes (34, Table 2.3). These included two genes categorized by MIPS to be involved in cell rescue, defense and virulence, namely *FGSG_16080* that encoded for a chitinase, and *FGSG_00132* with potential kinase function. The kinase *FGSG_00132* was previously targeted by reverse genetics, but no phenotype was observed in the mutant (Wang *et al.* 2011b). Another gene *FGSG_08005* that was not present in the NRRL28720 genome was annotated to be involved in isoprenoid metabolism, potentially contributing to secondary metabolic pathways (Accompanying Materials: Table S2.6).

Our analysis also showed that some of the genes in the accessory genome had irregular patterns of gene conservation across the ten genomes; many of these genes were associated with secondary metabolism (Table 2.4). An example includes *FGSG_04694*, a gene that encodes a polyketide synthase (PKS2) that was absent in both *F. asiaticum* genomes and was present in *F.*

graminearum and *F. meridionale* genomes (Table 2.4). Mutation of *PKS2* in *F. graminearum* showed reduced mycelial growth with a defect in virulence (Gaffoor *et al.* 2005). Another polyketide synthase, *FGSG_06540* was absent in both *F. meridionale* genomes, and present in *F. graminearum* and *F. asiaticum* genomes (Sieber *et al.* 2014). Another example includes gene *g11843* from NRRL28723 (*F. meridionale*), which shares similarity with *FAVG1_07768*, a gene from *F. avenaceum* and was present in the genomes of both *F. meridionale* isolates and the *F. asiaticum* isolate NRRL6101, but is absent in the genomes of all *F. graminearum* isolates. This gene was annotated as a chalcone synthase, which was extensively characterized in plants and found to be involved in the biosynthesis of defensive flavonoid compounds (Dao *et al.* 2011). In addition to polyketide synthases, a metabolic gene *g8968* from NRRL28721 (*F. meridionale*) shares identity with a terpene synthase gene *Tri5*, which catalyzes the first step of trichothecene biosynthesis in *F. graminearum* and is essential for trichothecene biosynthesis and infection of wheat (Jansen *et al.* 2005). Interestingly, *g8968* did not have a positive BLASTn hit to the other *F. meridionale* genome NRRL28723 but had positive BLASTn hits to the *F. asiaticum* genomes NRRL6101 and NRRL28720. Upon closer examination, *g8968* showed only 21.8% protein identity to *Tri5* from the *F. graminearum* isolate PH-1, but 95% protein identity to *FAVG1_13210* from *F. avenaceum*; the genome of *F. avenaceum* is the focus of the study described in Chapter 5 (Lysøe *et al.* 2014). A true homologue of *Tri5* (g5480) with >97% protein identity was identified in the *F. meridionale* isolate NRRL28721. Despite the low identity to *Tri5*, *g8968* was predicted to contain a *Tri5* domain by Interpro and Pfam (e-value 5.3E-15). It is unclear if the protein from this gene acts on substrates from the trichothecene biosynthetic pathway and contributes to their production, or if the protein is involved in catalysis of other substrates.

In addition to including specific genes involved in secondary metabolism, accessory genes with irregular patterns of conservation were also observed in gene clusters. For example, a gene cluster composed of five genes, including *FGSG_08181*, was absent in both *F. meridionale* genomes, but present in the genomes of *F. graminearum* and *F. asiaticum* isolates (Table 2.4) (Geiser *et al.* 2013). The genes within the *FGSG_08181* cluster include *FGSG_08182*, which encodes a transcription factor, and three genes *FGSG_17088*, *FGSG_08183*, and *FGSG_08187*, that encode cytochrome P450s (Geiser *et al.* 2013). Genes from this cluster have been reported to be expressed during infection of different cereal hosts (Harris *et al.* 2016). Similarly, another gene cluster (*FG05_30489* - *FG05_30492*) from the *F. graminearum* isolate CS3005 from Australia was absent in the 15-ADON producing isolates of *F. graminearum* from North America, but was present in all other genomes (Table 2.4). The gene *FG05_30489* with similarity to a polyketide synthase gene *PKS43* is also present in *F. avenaceum* and *F. equiseti* genomes (Hansen *et al.* 2015).

In addition to the above mentioned accessory genes potentially involved in secondary metabolism, there were other genes from the accessory genome with irregular conservation patterns, some of which have been previously characterized by gene disruption (Table 2.5). A study which globally targeted transcription factors determined that accessory genes *FGSG_08182* and *FGSG_11561* contributed to disease and spore production respectively (Son *et al.* 2011). A separate study which focused on kinases determined that the accessory genes *FGSG_07812*, *FGSG_11614*, and *FGSG_02153* all contributed to toxin production while *FGSG_03146* contributes to spore production and disease (Wang *et al.* 2011b). Disruption of a possible phosphatase regulatory subunit, *FGSG_11641* has also been suggested to reduce disease accumulation (Yun *et al.* 2015). Together, the accessory genome appears to contain genes that

contribute to metabolic processes, as well as other processes potentially involved in development or disease.

2.4 Discussion and Conclusions

Fusarium Head Blight fungi are globally distributed and isolates are known to exhibit phenotypic diversity in their morphology, growth and development, production of secondary metabolites and toxins, and pathogenicity towards specific hosts (O'Donnell *et al.* 2000). Phylogenetics using gene markers have allowed us to identify distinct species and genetic populations of these fungi; however, some isolates have proven difficult to classify using marker analyses and the biological mechanisms underlying phenotypic diversity remain unsettled (Carter *et al.* 2002; O'Donnell *et al.* 2000; O'Donnell *et al.* 2004; Starkey *et al.* 2007; Ward *et al.* 2002; Yang *et al.* 2008). Genome sequencing of additional isolates provides the opportunity to better understand and confirm the genetic relationship between fungal isolates, and provides gene candidates that could contribute to phenotypic differences and niche adaptation. Similar approaches have been successfully applied to the *F. oxysporum* species complex; for example, genomic comparison of *F. oxysporum* isolates identified genes that may be associated with pathogenicity towards banana, as well as candidate avirulence genes associated with infection of melon (Guo *et al.* 2014; Schmidt *et al.* 2016).

A comparative genomics approach used in this study was effective at better resolving the genetic relationship between fungal species and isolates. Genome alignments of the ten closely related fungi showed greater alignment to each other, particularly within species; the genomes also had greater alignment to *F. pseudograminearum* and *F. langsethiae*, than to genomes from other fusaria, or other genera altogether. These findings confirm the previously established phylogenetic relationships between these closely related *Fusarium* species to each other and

other fungi (O'Donnell *et al.* 2000; Ma *et al.* 2010; Starkey *et al.* 2007). Interestingly, on average, *F. meridionale* and *F. graminearum* shared more genes and had greater genome alignments than *F. asiaticum* and *F. graminearum*. Based on these findings, *F. meridionale* and *F. graminearum* may be more closely related than *F. asiaticum* and *F. graminearum*. In addition to providing insights into the relationships between species, we were also able to better resolve species placement of NRRL28721, which was difficult to determine using marker analyses. Marker analyses suggested that NRRL28721 is either *F. asiaticum* or *F. meridionale* or a species hybrid (Carter *et al.* 2002; O'Donnell *et al.* 2000; O'Donnell *et al.* 2004; Starkey *et al.* 2007; Ward *et al.* 2002; Yang *et al.* 2008). Our analyses indicate a 99.0% alignment between the genomes of NRRL28721 and NRRL28723; the analyses of SNPs and indels across genomes also confirm that these two isolates are highly related and likely of the same species, *F. meridionale*. We were also able to use SNP and indel information between genomes to separate the genetic populations that have been previously reported to exist within *F. graminearum*, which also correlated with the production of either 3-ADON or 15-ADON (Liang *et al.* 2015). Interestingly, CS3005, which is a 15-ADON producing isolate from Australia falls within the same population as the 15-ADON producers, indicating that it may be more closely related to the 15-ADON producers from North America than the 3-ADON producers from North America. This suggests that genetic populations are globally distributed and that geographical barriers may not play a significant role in the occurrence of these fungi. Although *F. graminearum* is homothallic, recombination rates (crossovers/bp/generation) from 0.29E-5 to 7.28E-5 have been reported; recombination may affect the species placement in our genome alignment analyses (Talas *et al.* 2015).

The presence of different species, such as *asiaticum* and *meridionale* in Nepal, or genetic populations of the same species, such as 3-ADON and 15-ADON populations of *F. graminearum* in North America, that cause similar diseases indicates that different populations and species can coexist and may have distinguishable traits that allow them to occupy slightly different niches than other populations or species. For example, environmental and host conditions have been demonstrated to affect disease outcomes by 3-ADON and 15-ADON producers of *F. graminearum* (Gilbert *et al.* 2014). Further monitoring of these populations and identification of undiscovered populations will better resolve the geographical distribution of these populations/species and the stability of the population and community structure over time; this may be particularly important when we consider our intensive agriculture practices and the environmental effects of climate change.

In addition to providing insights into the genomic relationship between isolates, we also established patterns in gene content. We observed a high degree of genetic overlap (95% to 99%) between the ten genomes, which was higher within species than between species (Figure 2.6). The gene content analyses of the two 3-ADON producing strains of *F. graminearum* (NRRL28336 and DAOM241165) showed the fewest number of absent genes (914 and 990, respectively). In contrast, the two *F. asiaticum* isolates NRRLs 6101 and 28720 had the greatest number of absent genes, 1126 and 1146, respectively (Accompanying Materials: Table S2.2). Further study across additional isolates would be required to determine if there is gene expansion or loss in these populations or species, respectively.

Many of the genes from the accessory genome were discovered to be involved in metabolism. This is particularly important because *Fusarium* pathogens are well known for their ability to make a wide array of secondary metabolites, including the terpene deoxynivalenol,

which has been studied for its role as a virulence factor during plant infection and for its toxic effects in humans (Jansen *et al.* 2005). Based on the sequence, *F. graminearum* isolate PH-1 is reported to have at least 16 known polyketide synthase (PKS) genes and 19 known non-ribosomal peptide synthases (NRPS) (Hansen *et al.* 2015); additional metabolic genes have been proposed in other studies (Geiser *et al.* 2013). We discovered that some of these genes, such as PKS2, PKS43 and NRPS17 that have been described previously in *F. graminearum*, are absent in some of the ten genomes (Hansen *et al.* 2015). Furthermore, metabolic genes from other *Fusarium* species may be present in some but not all of the ten genomes, including PKS40, PKS52, PKS43, and NRPS31 (APS1) (Hansen *et al.* 2015). We also identified genes or alleles from the accessory genome that have roles outside of secondary metabolism that may be implicated in disease; including kinases and transcription factors that have already been targeted in other studies (Table 2.5). The underlying mechanism by which these accessory genes are contributing to evolution and niche adaptation requires further investigation.

Many of the accessory genes, including those that we identified to be species or isolate specific, were annotated by Blast-2-Go based on sequence similarity to genes from other fusaria (Figure 2.7). There were 1061 gene sequences that we annotated from the new genomes (excluding CS3005 and PH-1) and 361 or 34% of these had a top tBLASTx hit in *F. oxysporum*, while the next top hits were in *F. pseudograminearum* (15%), *F. graminearum* (11%), *F. avenaceum* (7%), *F. fujikuroi* (5%), *Nectria haematococca* or *F. solani* (3%), and *F. verticillioides* (2%) (Figure 2.7). All other genus/species that had more than 0.5% of genes BLASTx hits were from other fusaria or other Ascomycete genera, including *Colletotrichum*, *Stachybotrys*, *Bipolaris*, and *Pseudogymnoascus*. This indicates that many genes from the accessory genome are present within the genus, and may be lost or acquired, depending on

selective pressures from the environment or plant host. This has been previously described as an attribute of pathogen effectors that are involved in the evolutionary interactions between pathogen and host.

At the same time, horizontal gene transfer may also play an important role in the biology of *Fusarium* species. This is supported by our observation that many of the genes that were absent across genomes were co-localized and may occur in gene clusters in regions of the genome with increased nucleotide variability (Figure 2.3). The co-localization of variable genomic regions and absent genes between related genomes is consistent with what has been reported previously both between and within *Fusarium* species (Ma *et al.* 2010; Gardiner *et al.* 2012; Walkowiak *et al.* 2015). Co-localization may suggest that genes were acquired or lost as clusters; multiple gains or losses of gene clusters have been suggested for the secondary metabolite gene cluster responsible for fumonisin biosynthesis in *Fusarium* (Proctor *et al.* 2013). In *F. oxysporum*, lineage-specific chromosomes have been observed to confer pathogenicity, and it has been observed that these chromosomes can be transferred horizontally to non-pathogenic strains causing them to become pathogenic (Ma *et al.* 2010). Lineage-specific chromosomes and largescale horizontal gene transfer events, such as the transfer of an entire chromosome that have been reported in *F. oxysporum*, are yet to be reported for members of the *F. graminearum* species complex. However, small scale horizontal transfer events have been suggested between *F. graminearum* and other fungi (Sieber *et al.* 2014). Horizontal gene transfer has also been suggested between bacteria and *F. pseudograminearum*, which is closely related to *F. graminearum* and is reported to cause disease on the roots of cereals (Gardiner *et al.* 2012). Closer inspection of the presence, location, and orientation within the genomes of other organisms could shed light on their history and origin.

In addition to accessory genes contributing to genetic differences amongst the genomes, we also identified common genes that are variable and thus could contribute to niche adaptation and disease. Common genes involved in plant-pathogen interaction that are recognized by the host are selected against but cannot be shed; instead, these genes modify their sequence in order to avoid recognition by the host (Jones and Dangl 2006). Our analyses identified 163 common genes with greater than 20% sequence variability. This included *Tri8*, a gene that is known to have allelic variation between fungal isolates and is involved in the production of trichothecene variation resulting in either 15-ADON or 3-ADON (Alexander *et al.* 2011). It is quite possible that many of the other rapidly evolving common genes are contributing to toxin production or disease in a similar manner, though most are uncharacterized. The potential role of these genes in niche adaptation or disease remains to be established.

To summarize, comparative genomics was able to separate the isolates into their species and genetic population. Although most genes were shared between genomes, we were able to identify accessory genes that were absent and common genes that were variable between genomes; these included those with functions that are associated with secondary metabolism and disease. Several of the metabolic genes that we identified to be variable have been reported in other *Fusarium* species, but are yet to be described in *F. graminearum* or its close relatives. As sequencing technologies become more accessible, larger comparative studies will allow for the further characterization of the genomic and genetic differences amongst these economically important pathogenic fungi.

2.5 Methods

2.5.1 *Fusarium* isolates, sequencing, and genome assembly

Total DNA was isolated from *Fusarium* isolates NRRL6101, NRRL28336, NRRL28720, NRRL28721, NRRL28723, DAOM233423, DAOM241165 and DAOM180378 using the E.Z.N.A. Fungal DNA Mini Kit (Omega). NRRL accessions are from the United States Department of Agriculture's Agriculture Research Service Culture Collection, while DAOM accessions are from the National Mycological Herbarium of Canada Department of Agriculture, Ottawa, Mycology collection. DNA was sequenced according to manufacturer instructions for 74 base pair (bp) reads from paired-end libraries by Illumina GAII sequencing for all isolates except DAOM233423 and DAOM241165, which were instead sequenced by using both mate-paired and paired-end libraries and resulted in 35bp reads. All sequencing reactions were performed by the Center for the Analysis of Genome Evolution and Function (CAGEF) at the University of Toronto, Canada. Sequencing reads were imported, processed, and *de novo* assembled by CLC Genomics Workbench, and contigs ordered according to the PH-1 reference (MIPS database v3.2). Prior to assembly, reads were trimmed and short reads <20bp were removed. Assemblies were optimized for contig length by adjusting the 'word size' parameter and keeping all other parameters at default settings. Average contig coverage was determined for each genome assembly and contigs that were 75% or lower than the average contig coverage and contigs shorter than 200bp were removed from the assembly. Contigs were then reordered to the genome sequence of PH-1 from the MIPS database v3.2, and genomes were using ABACAS (Assefa *et al.* 2009). The ordered contig sequences were uploaded to the GenBank database at NCBI (Table 2.2).

2.5.2 Sequence analyses, comparisons, and annotation

The newly sequenced genomes, and the genomes for PH-1 (MIPS v3.2) and for CS3005 that were obtained from GenBank at NCBI, were all used in separate pairwise genome alignments using the NUCMER and PROMER algorithms in the software package MUMMER. NUCMER is designed for aligning highly conserved regions of DNA, while PROMER aligns more divergent sequences. Since all matching and alignment routines are performed on the six frame amino acid translation of the DNA input sequence, PROMER is more sensitive than NUCMER. NUCMER was also used to extract the SNP and indel data, which was obtained by default parameters of the 1-to-1 algorithm in dnadiff (Delcher *et al.* 1999; Kurtz *et al.* 2004).

SNPs within conserved regions of the genome were obtained from PANSEQ with the following parameters: run mode = pan, fragment size = 500, percent identity cut-off = 90 and shared genome threshold = 10 (i.e. sequences were present in all genomes) (Laing *et al.* 2010). The resultant SNP alignment was imported into CLC Genomics Workbench for phylogenetic analyses using the maximum likelihood phylogeny tool, and the phylogeny was constructed by the neighbor joining method and Jukes Cantor substitution model using 100 replicates; the bootstrap values at all nodes were 100.

New genes were predicted by AUGUSTUS (Stanke and Waack 2003). For the construction of the pan gene list, genes were compared to the gene sequences from PH-1 by BLASTn and genes were added to the gene list if the lowest e-value $> 1E-10$; the percent genome alignment to PH-1 by MUMMER was used to select the order in which genomes were selected to contribute to the pan gene list going from greatest percent alignment to lowest. All genes in the pan gene list not originating from PH-1 were annotated by Blast-2-Go (Conesa *et al.*

2005). Genes were then compared to the genome sequences and were considered absent if they had a lowest BLASTn e-value $>1E-10$. For the analysis of sequence variation, the number of positions that contained SNPs and indels from the dnadiff analysis in MUMMER were tabulated for the full length of the gene sequence for each gene from the pan gene list using the native genome as a reference. The number of positions in the gene containing variation across all genome alignments to the nine other genomes was then converted into a ratio per kilobase of sequence.

Figure 2.1 Concordance between genome alignments and taxonomy. Whole genomes were aligned between ten closely related isolates of *F. graminearum*, *F. asiaticum*, and *F. meridionale*, as well as 11 other fusaria, individuals from 21 different *Ascomycete* genera, and individuals from 11 different *Basidiomycete* genera. Alignments were performed using two different algorithms, a more stringent aligner, ‘NUCMER’ (left), and a more relaxed aligner ‘PROMER’ (right). Genomes are arranged by species; *F. graminearum* (FG), *F. asiaticum* (FA), and *F. meridionale* (FM). A non-linear full spectrum heat map is used to represent the percentage alignment to each query genome, where red has the lowest alignment and violet has the greatest alignment.

Figure 2.2 Phylogeny of FHB pathogens by whole genome alignment. A SNP and indel alignment of common genomic regions was performed by PANSEQ and was used to construct a maximum likelihood phylogeny by neighbor joining using 100 replicates. All nodes have bootstrap values of 100.

Figure 2.3 Sequence and gene content variability between genomes. SNPs/indels and absent genes between genomes were determined separately, by whole genome alignments and BLASTn, using each genome as a reference. **A**, *F. graminearum* genomes were used as a reference. **B**, *F. meridionale* genomes are used as a reference. **C**, *F. asiaticum* genomes are used as a reference. The number of SNPs and indels per 100kb between the reference genome and the average for each species is presented separately for *F. graminearum* (black lines), *F. asiaticum* (red lines), and *F. meridionale* (blue lines). The minimum and maximum for each plot is 0 and 10000 nucleotide positions. Positions of absent genes whose sequences had a lowest expected value $> 1E-10$ by BLASTn for at least one member of the other species are presented as symbols; *F. graminearum* (black diamonds), *F. asiaticum* (red diamonds), and *F. meridionale* (blue diamonds). Absent genes in all other members of all species were considered unique genes (purple squares). Some regions of SNPs /indels that are greater in a specific reference genome than in other reference genomes are highlighted (dotted box).

Figure 2.4 The accessory genome is a subset of the pan genome. The pan genome, which incorporates the genes from all ten genomes, is composed of 15297 genes, 1827 of which are accessory genes that are absent in at least one genome. The accessory genome further divides into genes that are specific to each species, with 105 genes specific to *F. graminearum* (FG), 126 specific to *F. meridionale* (FM), and 184 specific to *F. asiaticum* (FA).

Figure 2.5 Sequence variability within common genes. Only 163 genes contain greater than 20% sequence variation. Genetic variability of genes from the common genome were assessed based on SNPs and indels in those genes from the genome alignments.

Figure 2.6 Percent gene overlap between genomes. The percentage of genes from the query genome in the reference genome are presented in a greyscale heatmap, where white represents 95% overlap and black represents 100% overlap. Genes are considered present if they have a lowest BLASTn expected value $<1E-10$. Genomes are arranged by species; *F. graminearum* (FG), *F. asiaticum* (FA), and *F. meridionale* (FM).

Figure 2.7 Top BLASTx hit of genes from the pan gene list not from PH-1 or CS3005. The top 13 genus/species matching to genes are presented in descending order and are presented as a percentage of genes; all other species/genus had less than five genes (0.5% of genes) that were a top hit to that genus/species.

Table 2.1 Description and Origin of *Fusarium* Isolates Used in this Study

Isolate Accession	Species	Disease on Wheat ^{a, b}	Trichothecene Toxin ^b	Origin	
				Host	Location
PH-1	<i>F. graminearum</i>	High	15-ADON, DON	Maize	USA
CS3005	<i>F. graminearum</i>	High	13-ADON, DON	Barley	Australia
DAOM233423 (GZ3639)	<i>F. graminearum</i>	High	15-ADON, DON	Wheat	USA
DAOM241165	<i>F. graminearum</i>	High	3-ADON, DON	Wheat	Canada
DAOM180378	<i>F. graminearum</i>	High	15-ADON, DON	Maize	Canada
NRRL28336	<i>F. graminearum</i>	Medium	3-ADON, DON	Wheat	USA
NRRL6101	<i>F. asiaticum</i>	None	3-ADON, DON	Barley	Japan
NRRL28720	<i>F. asiaticum</i>	Medium	DON	Maize	Nepal
NRRL28721	<i>F. meridionale</i>	Low	4A-NIV, NIV	Maize	Nepal
NRRL28723	<i>F. meridionale</i>	Low	NIV	Maize	Nepal

^a High = more than six florets infected, Medium = three to six florets infected, Low = infection on florets adjacent to inoculation site, None = no disease beyond the inoculation site

^b Toxin analyses and pathology testing were performed in culture and/or on wheat (Alexander *et al.* 2011; Gardiner *et al.* 2009; Gilbert *et al.* 2002; Miller *et al.* 1983; O'Donnell *et al.* 2000; Vesonder *et al.* 1982; Ward *et al.* 2008)

Table 2.2 Genome and Assembly Statistics of *Fusarium* Isolates Used in this Study

Isolate	Genome Size (Mb)	N50	Coverage (fold)	NCBI ^a Accession	Genome Reference
PH-1	36.6	5.37	10	AACM	Cuomo <i>et al.</i> 2007
CS3005	36.7	0.46	40	JATU	Gardiner <i>et al.</i> 2014
DAOM233423	36.5	0.17	86	LAJZ0	Walkowiak <i>et al.</i> 2015
DAOM241165	36.6	0.44	144	LAKA	Walkowiak <i>et al.</i> 2015
DAOM180378	36.4	0.16	53	LHUC	This study
NRRL28336	36.7	0.40	36	LHUD	This study
NRRL6101	36.5	1.21	111	LHTY	This study
NRRL28720	36.4	0.35	36	LHTZ	This study
NRRL28721	36.5	0.11	51	LHUA	This study
NRRL28723	36.4	0.41	32	LHUB	This study

^a National Center for Biotechnology Information

Table 2.3 Isolate-Specific Genes Identified from Accessory Genome

Isolate	Uniquely Present Genes ^a	Uniquely Absent Genes ^b
PH-1	14	13
DAOM180378	18	8
NRRL28336	92	2
DAOM233423	18	10
DAOM241165	12	1
CS3005	14	6
NRRL6101	65	34
NRRL28720	72	29
NRRL28721	18	2
NRRL28723	22	7

^a Uniquely present genes were considered present in the genome of the isolate and absent in genomes of all other isolates; absence was determined by having a BLASTn hit with the lowest expected value $>1E-10$ in the genome of the isolate

^b Uniquely absent genes were considered absent in the genome of the isolate and present in the genomes of all other isolates; absence was determined by having a BLASTn hit with the lowest expected value $>1E-10$ in the genome of the isolate

Table 2.4 Secondary Metabolic Genes are Part of Accessory Genome

Gene Name	Source	Possible Function	Alternate Name(s)	Presence ^a
Polyketide metabolism				
<i>FGSG_04694</i>	PH-1	polyketide synthase	PKS2	FG, FM
<i>FGSG_06540</i>	PH-1	polyketide cyclase		FG, FA
<i>g11843</i>	NRRL28723	chalcone synthase b		FM, FA
<i>g6531</i>	NRRL6101	polyketide enoyl-reductase		FA
<i>g1989</i>	NRRL6101	polyketide synthase	PKS40	FA
<i>g21</i>	NRRL28723	polyketide cyclase dehydrase protein		FM
<i>FG05_30424</i>	CS3005	polyketide synthase	PKS52	FG, FA
<i>FG05_30565</i>	CS3005	snoal-like polyketide cyclase family protein		FG, FM
<i>FG05_30492</i>	CS3005	polyketide synthase		FG, FM, FA
<i>FG05_30491</i>	CS3005	polyketide synthase		FG, FM, FA
<i>FG05_30490</i>	CS3005	polyketide synthase		FG, FM, FA
<i>FG05_30489</i>	CS3005	polyketide synthase	PKS43	FG, FM, FA
Non-ribosomal peptide metabolism				
<i>FGSG_10702</i>	PH-1	non-ribosomal peptide synthase	NRPS17	FG
<i>g11745</i>	NRRL6101	non-ribosomal peptide synthase		FA
<i>g6644</i>	NRRL6101	non-ribosomal peptide synthase	APS1, NRPS31	FA
Terpene metabolism				
<i>FGSG_08181</i>	PH-1	terpene synthase		FG, FA
<i>g8968</i>	NRRL28721	Tri5-like terpene synthase		FM, FA

^a Gene was present in the genome of at least one isolate of *F. graminearum* (FG), *F. asiaticum* (FA), or *F. meridionale* (FM)

Table 2.5 Accessory Genes Contribute to Virulence, Toxin Production, and Development

Gene	Mutant Phenotype	Presence	Reference
<i>FGSG_08182</i>	Reduced Virulence	FG, FA	Son <i>et al.</i> 2011
<i>FGSG_11561</i>	Reduced Conidiation	FG, FA	Son <i>et al.</i> 2011
<i>FGSG_11641</i>	Reduced Virulence	FG	Yun <i>et al.</i> 2015
<i>FGSG_03146</i>	Reduced Conidiation and Virulence	FG, FA	Wang <i>et al.</i> 2011b
<i>FGSG_07812</i>	Reduced Toxin Production	FG	Wang <i>et al.</i> 2011b
<i>FGSG_11614</i>	Reduced Toxin Production	FG	Wang <i>et al.</i> 2011b
<i>FGSG_02153</i>	Reduced Toxin Production	FG	Wang <i>et al.</i> 2011b

^a Gene was present in the genome of at least one isolate of *F. graminearum* (FG), *F. asiaticum* (FA), or *F. meridionale* (FM)

Chapter 3: Intraspecies Interaction between Two *F. graminearum* Isolates Leads to Reduced Toxin Production and Disease

3.1 Abstract

FHB fungi coexist and interact with several different fusaria, as well as other plant pathogenic fungi and bacteria in the field. In Canada, *F. graminearum* exists as two main trichothecene chemotypes: 3-acetyldeoxynivalenol (3-ADON) and 15-acetyldeoxynivalenol (15-ADON). To understand the potential interactions between two isolates of these toxin chemotypes, we conducted co-inoculation studies both in culture and *in planta*. The studies showed that intraspecies interaction reduces trichothecene yield in culture and disease symptoms in wheat. To elucidate the genes involved in the intraspecies interaction, expression profiling was performed on RNA samples isolated from co-inoculated cultures and potential genes were identified by using the genome sequences of the respective isolates.

3.2 Introduction

As discussed in Chapter 2, FHB fungi have distinct physical, physiological, and molecular profiles, which are likely attributed to clear differences in their genomic and genetic makeup (Doohan *et al.* 1998; O'Donnell *et al.* 2004; Geiser *et al.* 2013; Wagacha *et al.* 2012). A survey of *Fusarium* disease in South Africa and a separate survey in Korea suggested that different species coexist in the same geographic regions and are better adapted to cause infection within specific hosts (Boutigny *et al.* 2011; Lee *et al.* 2009). It has also been suggested that different fusaria interact, either directly or indirectly, with each other and other microbes in the field; these interactions can be antagonistic or synergistic (Doohan *et al.* 1998; Xu *et al.* 2005). Interspecies interactions between fusaria *in vitro* have been demonstrated to affect germination

and germ tube growth; species also have zones of inhibition between them, indicating individuals can detect and respond to each other (Wagacha *et al.* 2012). Field trials with *F. graminearum* and *F. culmorum* isolates inoculated either individually or together suggested that isolates can outcompete one another and the competition outcome was dependent on various environmental factors (Von der Ohe and Miedaner 2011). The importance of environment is underscored by another co-inoculation study, where co-inoculation with *F. graminearum* and *F. verticillioides* resulted in either increased or decreased abundance of a single species depending on the field conditions (Picot *et al.* 2012). Interestingly, the authors also reported natural contamination from other *Fusarium* spp., thus underscoring the complex interactions that exist between multiple microbes in the field (Picot *et al.* 2012). Although numerous studies showed fusaria coexist within similar geographic areas and likely interact with each other and other microbes, little is known about direct intraspecies interactions.

In Canada, FHB is predominantly caused by the species *F. graminearum*. The two chemotypes that dominate Canadian farms are 3-acetyldeoxynivalenol (3-ADON) and 15-acetyldeoxynivalenol (15-ADON). Genetic markers have been able to identify at least two distinct genetic populations of *F. graminearum* in North America, which largely separate based on these chemotype distinctions (Liang *et al.* 2015). Allele differences in the biosynthetic gene *Tri8* have been attributed to the difference in chemotype (Alexander *et al.* 2011). A study suggested that 3-ADON isolates are more toxigenic and have increased growth capacity compared to the 15-ADON producing isolates (Ward *et al.* 2008). The same study also indicated that the distribution and prevalence of these chemotypes across the country differs and is likely dynamic, with 3-ADON potentially displacing the 15-ADON chemotype (Ward *et al.* 2008).

Chemotype prevalence has also been shown to be variable and linked to several different factors, such as temperature and wheat cultivars (Foroud *et al.* 2012; Gilbert *et al.* 2014).

We sought to investigate interactions between 3-ADON (FG1 or DAOM241165) and 15-ADON (FG2 or DAOM233423) *F. graminearum* isolates at the molecular level. The two isolates were assessed for their toxin production in culture and pathogenicity in wheat, independently and in combination. The data suggested that the two isolates are capable of producing toxins and are virulent; however, in combination both the trichothecene production and disease symptoms were reduced. The genomes of these two *F. graminearum* isolates have been described in Chapter 2, as such; we were able to exploit genetic differences between the isolates to monitor their abundance in mixed cultures. RNA-sequencing (RNA-Seq) also identified genes that behaved differently between isolates and genes that are responsive to intraspecies interaction; this included genes that were absent between the isolates.

3.3 Results

3.3.1 Interference between *F. graminearum* isolates affects trichothecene production *in vitro*

In Canada, FHB is primarily caused by the species *F. graminearum* and the dominant chemotypes are the 3-ADON or 15-ADON trichothecenes (Ward *et al.* 2008). Since these chemotypes coexist in the same geographic regions, we sought to dissect possible interactions between isolates of opposing chemotype. To assess direct interactions between isolates affecting toxin production, we used a modified two stage media method to induce trichothecene production. Toxin-type for each isolate was determined by high-performance liquid chromatography (HPLC) and nuclear magnetic resonance (NMR). HPLC was able to resolve 3-ADON in the FG1 culture supernatant at ~13.5min and 15-ADON at ~15min in the FG2 culture

supernatant; the opposing chemotype was not detected for either isolate (Figure 3.1). NMR analysis of the ethyl acetate extracts of the culture filtrates from the two isolates also showed the presence of 3-ADON for FG1 and 15-ADON for FG2 (Savard and Blackwell 1994; Walkowiak *et al.* 2015).

Prior to investigating if interactions between FG1 and FG2 affect trichothecene production, we assessed whether or not reducing the starting inoculum affected mycotoxin production. The results indicated that starting the inoculum at 5000 spores/mL and progressively diluting the inoculum to 90%, 50% and 10% with water did not significantly affect the production of either mycotoxin (grey bars, Figure 3.2 A and white bars, Figure 3.2 B). Moreover, regardless of the amount of starting inoculum, FG1 consistently produced half of the amount of ADON when compared to FG2 (compare grey bars in Figure 3.2 A to white bars in Figure 3.2 B). To determine if either of the isolates influenced mycotoxin production in the other, we quantified ADONs in FG1 and FG2 axenic cultures and compared it to co-inoculated cultures in a ratio of 10:1, 1:1, and 1:10, while maintaining the total concentration of spores at 5000 spores/mL. FG1 axenic culture produced 285µg of 3-ADON (FG1, Figure 3.2 A), but only produced 121µg when it was co-inoculated with the FG2 in a ratio of 10:1 of FG1:FG2 (striped bars, 10:1, Figure 3.2A) and 11µg when the culture was co-inoculated with equal amount of the FG2 isolate (striped bars, 1:1, Figure 3.2 A). In cultures where the FG2 isolate dominated, we could not detect any 3-ADON from FG1 (striped bars, 1:10, Figure 3.2A). We also quantified 15-ADON from FG2 axenic culture and compared it to co-inoculated cultures. FG2 axenic culture produced 665µg of toxin (FG2, Figure 3.2 B), 528µg when the culture was co-inoculated with FG1 isolate in a ratio of 10:1 of FG2:FG1 (striped bars, 10:1, Figure 3.2 B), and 370µg when the culture was co-inoculated with equal amount of the FG1 (striped bars, 1:1, Figure 3.2

B). In cultures where the FG1 isolate dominated, we detected 132 μ g 15-ADON (striped bars, 1:10, Figure 3.2 B). A quantitative real-time PCR (qRT-PCR) analysis was used to verify relative abundance of each isolate in the culture. A gene marker unique to each isolate (*g6419* for FG1 and *g6823* for FG2) showed that the abundance is similar to that of the initial inoculum, which suggested that the reduction in toxin production is not caused by an isolate competitively excluding growth of the other (Figure 3.2 C).

To further rule out the possibility that co-inoculation was influencing growth rates, we determined the growth rate of each isolate alone or in different ratios in both nutrient limiting minimal medium and in nutrient-rich potato dextrose medium by optical density in liquid and by radial growth on solid medium (Figure 3.3). Both isolates had similar growth rates in the nutrient limiting minimal medium, regardless of whether they were grown in liquid or solid (Figure 3.3 A and C). We observed differences in growth between isolates on the solid nutrient-rich potato dextrose medium. After 4 days, FG1 had increased (~1.5cm) radial growth on the solid medium compared to FG2 (Figure 3.3 D). In contrast, no differences were observed between the two isolates in the liquid potato dextrose medium (Figure 3.3 B). We also observed no significant changes in growth in either media conditions, when isolates were co-inoculated (Figure 3.3).

3.3.2 Interference between *F. graminearum* isolates affects disease outcome *in planta*

To determine if the decrease in toxin production of the two isolates in co-inoculated conditions influenced disease outcome, we point inoculated spores of the two isolates either alone or in combination on a susceptible wheat cultivar, ‘Roblin’, and assessed for disease symptoms and quantified DON in the infected wheat heads. After 21 days post inoculation (dpi), point inoculation of either isolate alone caused disease symptoms, with FG1 causing slightly

more disease symptoms (96% infection) than FG2 (83% infection), p -value < 0.01 (FG1 and FG2, Figure 3.4 A). We observed a direct relationship between increased disease symptoms and DON accumulation (481ppm in FG1 and 365ppm in FG2 (Figure 3.4 B)). Diluting the spores of either isolate with water in a ratio of 10:1, 1:1, or 1:10, thereby decreasing the inoculum from approximately 1000 spores to approximately 900, 500, or 100 spores, respectively, did not cause a statistically significant difference in the amount of disease symptoms for either isolate (FG1: H₂O and FG2:H₂O, Figure 3.4 A); however, when the inoculum was reduced, there was a reduction in the amount of DON that accumulated in the wheat heads (481ppm for 1000 spores of FG1 to 317ppm for 100 spores; and 365ppm for 1000 spores of FG2 to 216ppm for 100 spores (Figure 3.4 B)).

Next, we performed a co-inoculation experiment, while maintaining the total spore inoculum at 1000. As results indicate, co-inoculation with FG1 to FG2 in the ratio of 1:1 and 10:1 resulted in 88% and 92% infection, respectively (FG1:FG2, Figure 3.4 A), both lower than the sole inoculum of FG1, but greater than a sole inoculum of FG2 (FG1, FG2, Figure 3.4 A). Co-inoculation of FG1 and FG2 in the ratio of 1:10 resulted in significant reduction in disease symptoms (71% infection) (FG1:FG2, Figure 3.4 A), compared to corresponding dilution of FG1 with water (95% infection) or FG2 with water (87% infection), p -value < 0.05 (Figure 3.4 A). The reduction in disease symptoms in this condition was reflected with a reduction in DON accumulation, 150ppm, which is less than half of the equal inoculum of FG1 (317ppm) or FG2 (313ppm) (Figure 3.4 B).

Since we observed differences in virulence during co-inoculated conditions, we were also interested to know which *Fusarium* isolate dominated at the end point of infection, which was assessed by both PCR and qRT-PCR using isolate specific markers (*g6419* for FG1 and *g6823*

for FG2). qRT-PCR was used to detect the relative abundance of the isolates across several wheat heads (Figure 3.4 C), while PCR allowed us to determine the presence or absence of the isolate in a single wheat head (Figure 3.5). As expected, when the wheat heads were inoculated with one isolate, only the inoculated isolate was detected (Figure 3.4 C and Figure 3.5 A and E). When the composition of the inoculum contained an equal ratio of the two isolates, the wheat heads were mostly dominated by FG1 at the end of infection (Figure 3.4 C). When the wheat heads were inoculated with spores dominated by FG1, we only detected FG1 (Figure 3.4 C and Figure 3.5 B). However, when FG2 was the dominant inoculum, we detected both FG2 and FG1 in wheat heads albeit, FG2 was more abundant (Figure 3.4 C and Figure 3.5 D). In several instances, we were also able to detect both isolates in a single wheat head, particularly in the co-inoculated conditions where FG2 dominated the inoculum, which suggested that isolates are interacting throughout the infection process, including at the end point of 21 days (Figure 3.5).

Collectively, our results suggest that the two isolates infect with different degrees of severity and that combining isolates results in a reduction in disease symptoms. This was particularly evident in the co-inoculated conditions, when the FG1 and FG2 isolates were present in the ratio of 1:10. The same ratio also resulted in decreased mycotoxin production in culture (Figure 3.2) and in wheat (Figure 3.4 B), thus emphasizing a link between mycotoxin accumulation and disease spread.

3.3.3 Genome comparisons reveal differences in gene content between a 3-ADON and a 15-ADON producing isolate

A comparison of the genomes of FG1 and FG2 in Chapter 2 indicated that the genomes were very similar; ~97% of the two genomes aligned with one another (Figure 2.1). There were however, some regions of the two genomes that did not align and some differences in gene content (Figures 2.1 and 2.6). FG1 and FG2 were also reported to have 12 and 18 genes respectively that were unique based on BLASTn analyses to the genomes of the nine other genomes from Chapter 2 (Table 2.3). We performed a more thorough analysis of gene presence and absence between FG1 and FG2 using the complete list of AUGUSTUS predicted genes from each genome as well as the genes from the PH-1 MIPS database V3.2. We assessed gene presence and absence by BLASTn, as before, and also used RATT (Rapid Annotation Transfer Tool) which considers sequence alignment and position within the genome (Accompanying Materials: Table S3.1) (Otto *et al.* 2011). Included in the list of absent genes between genomes were gene clusters, such as the genes spanning *g8800* to *g8808* on contig_73, and genes spanning *g3545* to *g3564* on contig_207 in the FG1, which were absent in FG2; the latter of which is the focus of Chapter 4 (Accompanying Materials: Table S3.1). Specific genes that are absent in FG2 include *g3560* in contig_207 of FG1, which is predicted to encode a protease, and *g2283* on contig_127 of FG1, that is predicted to encode a beta-lactamase, which may be associated with antibiotic resistance (Accompanying Materials: Table S2.6) (Allen *et al.* 2010). Genes associated with heterokaryon incompatibility that may have a role in non-self recognition were also identified as absent between genomes, which included *g6823* (contig_53) in FG2 isolate and *g6419* (contig_122) in isolate FG1 (Glass and Kaneko 2003). It is noteworthy to

indicate that *g6823* and *g6419* were used as genetic markers for each isolate and were used in both PCR and quantitative PCR analyses (Figures 3.2 C, 3.4 C and 3.5).

3.3.4 Transcriptome profiling of axenic FG1 and FG2 isolates in mycotoxin-inducing conditions

We performed RNA-Seq from axenic cultures grown in the toxin-inducing conditions from five independent biological replicates. Sequence reads were mapped separately to coding sequences of PH-1, FG1 and FG2 isolates using CLC Genomics Workbench. First, we compared the gene expression profiles of axenic cultures directly by DESeq. The profiles revealed that 1600-1800 or ~13-15% genes were found to be differentially expressed between FG1 and FG2 by two-fold or greater (p-value < 0.05) (Table 3.1). Differentially expressed genes included those involved in metabolism, toxin production, energy, and cell rescue/defence/virulence as well as cell type differentiation (Accompanying Materials: Table S3.2; Accompanying Materials: Table S3.3). Genes that were absent between the genomes were also found to be differentially expressed, further supporting their lack of conservation (Accompanying Materials: Table S3.3).

Specifically, the *Tri* genes, which are involved in trichothecene production, had greater relative expression in FG2 than FG1 (Table 3.2). The expression patterns determined by RNA-Seq were consistent, regardless of which genome was used as reference. Minor variations that were observed between references, for example in *Tri8*, may be attributed to allelic differences between the isolates (Table 3.2). We verified the RNA-Seq data with qRT-PCR of two genes, *Tri6* and *Tri5*. The relative quantification (RQ) values confirmed that *Tri6* had increased expression in FG2 by 4.2-fold, whereas, *Tri5* was increased in this isolate by 7.9-fold with respect to FG1. Despite the different levels of expression, both isolates are able to make their

respective trichothecene product, albeit, FG2 produces more than the FG1 (Figure 3.1 and 4.2). Interestingly, there was incongruity with respect to the expression of the two regulatory genes, *Tri6* and *Tri10*. Although *Tri6* was up-regulated in FG2, similar to other *Tri* genes, we observed no differences in the expression of *Tri10* between the two isolates (Table 3.2).

3.3.5 Transcriptome profiling of co-inoculated cultures in mycotoxin-inducing conditions

Since the co-inoculated samples contained RNA originating from both isolates, they could not be compared directly without accounting for the abundance of each isolate in the sample. To determine the abundance of each isolate, qRT-PCR was performed with DNA extracted from the co-inoculated cultures (Figure 3.2 C). By knowing abundance of each isolate in the sample and the expression in axenic cultures, we constructed an expected gene expression profile for each co-inoculated sample; an example calculation for this is found in the materials and methods. The expected expression profile was then compared to the actual expression profiles of the co-inoculated cultures to obtain relative expression by DESeq.

In contrast to axenic cultures, less than 1.5% of the *Fusarium* genes are affected in all the co-inoculated culture conditions (Table 3.1). As with the previous analysis, we used all three isolates as a reference and found that regardless of the reference, similar numbers of genes are differentially expressed (Figure 3.6; Accompanying Materials: Table S3.3). Although, there was up to eight genes common to two co-inoculation conditions (FG1:FG2 ratio of 1:1 and 1:10), only two genes were differentially expressed in all three co-inoculation conditions (*FGSG_09051* and *FGSG_16275*), both of which are annotated as hypothetical proteins (Figure 3.6; Accompanying Materials: Table S3.3). When the expression of these two genes was

compared in the axenic cultures, *FGSG_09051* had increased expression in FG1, while *FGSG_16275* has similar expression patterns between the isolates.

To further validate the differential expression of these genes in the co-inoculated conditions, we chose four genes based on differential expression and the presence/absence in the respective genome and subjected them to qRT-PCR analyses. The results showed that RQ values obtained from the qRT-PCR analysis for all four genes were in agreement with the expression pattern from the RNA-Seq analyses. This included genes common to both isolates (*FGSG_09051* and *FGSG_13870*) and genes that were absent in FG1 (*FGSG_04587*) or absent in FG2 (*FGSG_04760*) (Figure 3.7).

The RNA-Seq from the co-inoculated conditions also suggested a down-regulation trend with respect to the *Tri* genes. We observed a consistent pattern of down-regulation, particularly in the cultures where FG1 was dominant (10:1) (Figure 3.8). We investigated this further with qRT-PCR analysis of *Tri6* and *Tri5* genes. Since the genome analysis showed polymorphism in the *Tri5* gene, we designed isolate specific primers for *Tri5*. As indicated, using isolate specific primers, cultures in which the FG1 is dominant (10:1), *Tri5* is significantly down-regulated (Figure 3.7 B). Although showing a trend of down-regulation, expression of *Tri6* or *Tri5* in other co-inoculated cultures, did not surpass the 2-fold cut-off for differential expression (Figure 3.8).

3.4 Discussion and Conclusions

Several studies have indicated that microbial communities can be diverse in a given geographic area, including within a single genus such as *Fusarium* (Boutigny *et al.* 2011; Lee *et al.* 2009; Ward *et al.* 2008). Undoubtedly, selective pressures have led the fungi to adapt to these communities. *Fusarium* isolates have been reported to produce different toxins and infect hosts

with different levels of aggressiveness, indicating that isolates can be diverse, each with its own unique characteristics (O'Donnell *et al.* 2000; Starkey *et al.* 2007). In North America, at least two distinct genetic populations of the fungus have been identified, which largely separate into two trichothecene chemotypes, 3-ADON and 15-ADON (Liang *et al.* 2015). Emerging isolates of 15-ADON and 3-ADON chemotypes may be of particular concern because of their increased aggressiveness in wheat (Foroud *et al.* 2012; Ward *et al.* 2008). Of particular interest to Canadian agriculture is the possibility of an increased prevalence of 3-ADON producing isolates in regions where 15-ADON producing isolates were previously more dominant (Ward *et al.* 2008). Several studies that have investigated the pathology differences between 15-ADON and 3-ADON producing isolates suggest that, on average, 3-ADON isolates are more aggressive than the 15-ADON isolates (Foroud *et al.* 2012; Gilbert *et al.* 2014; Puri and Zhong 2010; Ward *et al.* 2008). Our data supports these previous studies and showed that isolate DAOM241165 (FG1), which produces 3-ADON, caused more disease than the DAOM233423 (FG2) isolate, which produces 15-ADON (Figure 3.4). Moreover, wheat heads co-inoculated with an equal ratio of FG1 and FG2 resulted in disease dominated by FG1 (Figures 3.4 and 3.5).

Comparative genomics from Chapter 2 has already provided us with insights into the diversity of FHB pathogens and has allowed us to establish the “common” versus “accessory” genomes. Several studies indicate that genes encoded in the accessory genome can be acquired by recombination and horizontal gene transfer events, even in filamentous fungi such as *Fusarium* (Baltrus *et al.* 2011; Gardiner *et al.* 2012; Ma *et al.* 2010). Evolution of the accessory genome provides pathogens with the capacity to increase their diversity in an ongoing evolutionary tug of war between pathogen and host. In order to understand differences that exist in gene content genome between FG1 and FG2, analyses were performed with BLAST and

RATT software. This allowed us to identify that contig_207 from FG1, which contains 20 genes, some of which are absent in FG2; this contig is the focus of study described in Chapter 4.

In addition to the differences in gene content, transcript levels also differed between isolates when grown *in vitro*. RNA-Seq analyses showed striking differences in gene expression between FG1 and FG2 (Table 3.1). More than 1500 genes, representing more than 13% of *F. graminearum* genes were differentially expressed between the two isolates. Although these two isolates are very similar (~97%) at the genome level, the differential expression patterns suggest that these isolates are distinct and adapt differently to their environments. For instance, all of the essential genes involved in trichothecene production, except for *Tri10*, were up-regulated in FG2 compared to FG1 (Table 3.2). This profile was consistent regardless of which genome was used as a reference, thus confirming the distinctiveness of the two isolates. The differences in gene expression profiles of the two regulatory genes *Tri6* and *Tri10*, which positively regulate production of DON, suggests that expression of *Tri* genes is more closely linked to the activation of *Tri6*.

In addition to behaving differently, we report the outcomes of isolate interactions both *in planta* and *in vitro*. Overall, our observations indicate that interactions are inhibitory and result in reductions in infection and toxin accumulation *in planta*, particularly when in a ratio of 1:10 of FG1:FG2. *In vitro* we also observed interference between isolates with respect to toxin production; when in an equal ratio, the 3-ADON level dropped to ~11µg from ~285µg produced by the pure inoculum (~95% drop), whereas, in the same co-culture, the level of 15-ADON was reduced only from 665µg to 370µg (~45% drop) (Figure 3.2). This observation is mirrored in the expression profile of *Tri* genes, where many of the structural genes were down-regulated (Figure 3.8). Despite the large differences in gene expression profiles between the isolates (Table 3.1),

we found relatively few gene candidates that are affected by isolate interference. Together our results provide insights into the intraspecies interactions that occur at the molecular level between distinct *F. graminearum* isolates, and their consequences on toxin production and pathology in wheat.

Our results provide insights into the intraspecies interactions that occur at the molecular level between distinct *F. graminearum* isolates, and their consequences on toxin production and pathology in wheat. The results indicate that interactions between isolates result in reduced toxin production and disease, and that the 3-ADON producing isolate FG1 is more competitive than the 15-ADON producing isolate FG2. Further investigation is required to dissect the precise mechanisms of interference; however, we do observe some differences in gene content between isolates as well as a large difference in gene expression patterns between axenic cultures of the two isolates. This indicates that these two isolates are unique in their behaviour at the molecular level.

3.5 Materials and Methods

3.5.1 Fungal strains and inoculum preparation

Two *F. graminearum* isolates were used in this study; a 3-ADON producing isolate (DAOM241165; NRRL37440) from Manitoba (Canada), which we refer to as FG1 in the manuscript, and a 15-ADON producing isolate (DAOM233423; GZ3639; NRRL29169) from Kansas (U.S.A), which we refer to as FG2 in the manuscript. Macroconidia were used as the inoculum for all experiments and were produced in a liquid carboxymethylcellulose medium (Schreiber *et al.* 2011).

3.5.2 Trichothecene production *in vitro*

Trichothecene production of FG1 and FG2 was assessed in triplicate by modifying the two-stage media method (Nasmith *et al.* 2011). A nylon net filter was added to the culture plates. This allowed for even growth and avoided mycelia clumps. The culture conditions included spores from each isolate inoculated individually, or in combinations of the two isolates (in ratios of 10:1, 1:1, or 1:10), or dilutions of the isolates with medium (in ratios of 10:1, 1:1, or 1:10). After 24h in the GYEP medium, the mycelia that remained adhered to the nylon filter was washed several times with water before the addition of second stage media amended with putrescine at 1g/L as a sole nitrogen source (Nasmith *et al.* 2011). After incubating for an additional 48h under the same conditions, the supernatant of the culture was removed to quantify trichothecenes. Trichothecene accumulation in the supernatant was quantified with an isocratic gradient containing 15% methanol and purified chemical standards by high-performance liquid chromatography (HPLC) (Nasmith *et al.* 2011).

3.5.3 Growth assays

Growth was assessed in a liquid minimal medium and potato dextrose broth by inoculating 800 spores in a 96 well plate containing 200µL of medium (Leslie and Summerell 2006; Walkowiak and Subramaniam 2014). The plate was incubated at 28°C and optical density measurements were taken at a wavelength of 620nm using a POLARstar OPTIMA plate reader (BMG LabTech); the data was analyzed using FLUOstar OPTIMA 1.32 software (BMG LabTech). Growth was also assessed on a solid minimal medium and potato dextrose agar in 8.5cm diameter petri dishes using 20000 spores inoculated onto sterile 0.6cm diameter filter

paper disks placed in the center of dish. All cultures were grown in triplicate using the same spore combinations as before and incubated at 28°C in darkness.

3.5.4 Pathology testing in wheat and PCR/qRT-PCR analyses of isolates at the endpoint of infection

The wheat cultivar ‘Roblin’ was grown under controlled environmental conditions in growth chambers until mid-anthesis as described previously (Schreiber *et al.* 2011). The plants were point inoculated between the palea and the lemma with a total of 1000 macroconidia from either FG1 or FG2, combinations of the two isolates (in ratios of 10:1, 1:1, or 1:10), or dilutions of the isolates in water to decrease the inoculum (in ratios of 10:1, 1:1, or 1:10). Inoculated plants were then returned to growth chambers and grown with regular misting for two days. This was repeated as three independent trials for the combinations with water and four independent trials for the combinations with both isolates. The results from all seven trials were pooled for comparative analyses. Wheat heads inoculated with the FG1, FG2, or a combination thereof in ratios of 10:1, 1:1 or 1:10 were examined 21 days post inoculation (dpi).

To ascertain the presence of an isolate in individual wheat heads at 21 dpi, ten wheat heads for each condition were individually ground in liquid N₂ and DNA extracted by E.Z.N.A. Fungal DNA extraction kit (Omega). To test for the presence or absence of the isolates in the wheat samples, 100ng of DNA was used in a 40 cycle PCR reaction using primers specific to both the FG1 (*g6419* F/R) and FG2 (*g6823* F/R) isolates. Similarly, to determine the amount of DON and the relative abundance of the isolates across a greater sample set, 32 wheat heads from each condition divided evenly across the seven trials were also ground in liquid N₂. To determine the relative abundance of the isolates in the samples, DNA was isolated from the tissues as

before and qRT-PCR analyses were performed (Walkowiak and Subramaniam 2014). The primers *g6419* qPCR F/R was used to quantify FG1 and *g6823* qPCR F/R to quantify FG2; *Tri6*, a gene common to both isolates, was used as an internal control (Walkowiak and Subramaniam 2014). A complete list of primers used can be found in Accompanying Material Table S3.4. The amount of DON in the tissues was also quantified by Enzyme-Linked Immunosorbent Assay (ELISA) by services at AAFC ORDC (Sinha *et al.* 1995).

3.5.5 Genome sequencing, assembly, alignment, and gene prediction

Both FG1 and FG2 were sequenced and genes were predicted as described in Chapter 2. The presence and absence of genes predicted by AUGUSTUS and genes from PH-1 (MIPS V3.2) were determined by BLASTn as described in Chapter 2 and were also transferred between genomes by RATT by default parameters (Otto *et al.* 2011; Stanke and Waack 2003).

3.5.6 RNA isolation and RNA sequencing

Liquid cultures were prepared in duplicate for toxin induction as before for FG1 and FG2 both as sole isolation and in combinations of 10:1, 1:1, and 1:10. After 6h of incubation in the second stage medium, mycelia were harvested and flash frozen. RNA was isolated by Trizol (Life Technologies) and further purified by the Invitrap (Invitex). RNA sequencing was performed according to manufacturer instructions using 145bp reads from paired-end libraries by Illumina Miseq sequencing. The sequencing reactions were performed by the CAGEF at the University of Toronto, Canada. Reads were then imported into 'CLC Genomics Workbench 6.0.1' and trimmed using the same method as the genome sequencing. Reads were mapped to the coding sequences of PH-1, FG1, and FG2 genomes with the RNA-Seq Analysis tool in CLC Genomics Workbench.

3.5.7 Normalization, validation, and analysis of RNA sequencing data

To identify genes that are differentially expressed in response to interactions between isolates in the co-inoculated cultures, we constructed an expected expression profile dataset for the co-inoculated samples. The expected dataset was constructed by combining read data from each biological replicate of the pure cultures. The contribution of each isolate to the expected dataset was determined by the abundance of each isolate present in the co-inoculated cultures by qRT-PCR of DNA isolated from each culture. qRT-PCR was performed on DNA with isolate-specific primers as previously mentioned. Sequences of primers are listed in Accompanying Materials Table S3.4.

An example of the expected read counts for the gene *Tri10* in the co-inoculated sample of 10:1 ratio of FG1:FG2 is calculated as follows: $A = B + C$; where, 'A' is the expected number of *Tri10* reads in co-inoculated culture, 'B' is the expected number of *Tri10* reads from FG1 in pure culture, and 'C' is the expected number of *Tri10* reads from FG2 in pure culture. B and C are determined by the number of *Tri10* reads in pure culture divided by the total reads in pure culture, multiplied by the abundance of the isolate in the sample determined by qRT-PCR, multiplied by the total reads in co-inoculated sample.

Therefore, expected read counts for the gene *Tri10* in the co-inoculated sample is equal to: $(622/3805928 \times 0.94 \times 4680464) + (884/3815592 \times 0.06 \times 4680464) = 784$. This was calculated for all the genes in the co-inoculated sample. The read count data for the expected and actual data sets were normalized and compared in 'R' by 'DESeq', which is part of the 'Bioconductor' package (Anders and Huber 2010). A complete list of read mappings, both measured and expected, as well as the normalized relative expression values can be found in Supplementary Table S3.

To confirm the differential expression of select genes when comparing FG1 and FG2, or the co-inoculated samples to the expected datasets, qRT-PCR analyses were performed using transcript specific primers for differentially expressed genes using complementary DNA (cDNA) converted RNA (Walkowiak and Subramaniam 2014). Isolate-specific primers were also used based on single nucleotide polymorphisms for the gene *Tri5*. A complete list of primers used can be found in Accompanying Materials Table S3.4.

Figure 3.1 HPLC chromatographs that show abundance of 3-ADON and 15-ADON in FG1 and FG2 in axenic and co-inoculated cultures.

Figure 3.2 Co-inoculation of *F. graminearum* isolates reduces trichothecene production *in vitro*. **A**, yield of 3-ADON toxin in a 4ml culture by HPLC resulting from dilutions of FG1 with dH₂O (grey) and with FG2 (striped), in a ratio of 10:1, 1:1 and 1:10, respectively. **B**, yield of 15-ADON toxin in a 4ml culture by HPLC resulting from dilutions of FG2 with dH₂O (white) and with FG1 (striped) in ratios of 10:1, 1:1 and 1:10, respectively. Results are representative of three independent samples; errors bars reflect standard error. **C**, Relative abundance of each isolate is maintained in culture during trichothecene production. Isolates were either inoculated alone, or in a ratio of 10:1, 1:1, or 1:10 (FG1:FG2). The relative abundance of FG1 (grey) and FG2 (white) was determined independently using qRT-PCR primers, specific to each single isolate. Results represent the average of five independent samples tested in triplicate, error bars reflect the maximum and minimum across the five independent samples.

Figure 3.3 Growth of FG1 and FG2 isolates in both solid and liquid media. FG1 (black bars) and FG2 (white bars) were inoculated either alone or in a ratio of 10:1, 1:1, and 10:1 respectively (grey scale). Eight hundred spores were inoculated into 200 μ l of **A**, minimal medium (MM) or **B**, potato dextrose broth (PDB) and grown in a 96-well plate reader system which took regular measurements of optical density ($\lambda=620\text{nm}$). Twenty thousand spores were inoculated onto a sterile disk in the center of **C**, solid minimal medium (MM) or **D**, potato dextrose medium (PDA). Error bars from both panels reflect the maximum and minimum values from three biological replicates.

Figure 3.4 Co-inoculation of *F. graminearum* isolates affects disease outcomes in wheat. **A**, Wheat heads were point inoculated with spores from FG1 (n=107) and FG2 isolates (n=106), where, 'n' represents the number of wheat heads. Inoculations were also performed with spores from the FG1 diluted with water in a ratio of 10:1(n=43), 1:1(n=43), and 1:10 (n=42) (grey) (FG1:H₂O). Similarly, spores from the FG2 isolate were inoculated after dilution with water in a ratio of 10:1 (n=43), 1:1(n=41), and 1:10 (n=42) (white) (FG2:H₂O). Disease symptoms were also assessed with combinations of spores from FG1 and FG2 in ratios of 10:1 (n=75), 1:1 (n=76), and 1:10 (n=76) (striped) (FG1:FG2). Results are presented as percent spikelets with disease symptoms at 21 dpi. Error bars reflect standard error, while 'a' and 'b' denote-values that are statistically different with a *P*-value of 0.1 or below compared to sole inoculation of FG1 and FG2 isolates, respectively, by one-tailed Student's t-test. A single asterisk (*) represents a significance of $0.05 \geq P\text{-value} \leq 0.01$, while two asterisks (**) represents a significance of $P\text{-value} \leq 0.01$. The 1:10 ratio in bold highlights the condition where the co-inoculated isolates significantly decrease disease outcome compared to the each isolate diluted with water in the same ratio. **B**, Thirty-two wheat heads from each condition were tested for DON content by ELISA. The 1:10 ratio in bold highlights the conditions where there was a decrease in DON accumulation compared to equivalent amounts of the isolates diluted with water. Error bars reflect the maximum and minimum across three independent replicates. **C**, Relative abundance of each isolate is not maintained in wheat at 21 dpi. The relative abundance of FG1 (grey) and FG2 (white) was determined independently across 32 wheat heads from each condition using qRT-PCR primers, specific to each single isolate.

Figure 3.5 The FG1 isolate is more prevalent than the FG2 isolate at 21 dpi. Detection of FG1 and FG2 in **A**, wheat heads inoculated with FG1 alone, **B**, wheat heads inoculated with FG1 and FG2 in 10:1 ratio, **C**, wheat heads inoculated with FG1 and FG2 in 1:1 ratio, **D**, wheat heads inoculated with FG1 and FG2 in 1:10 ratio, and **E**, wheat heads inoculated with FG2 alone. The presence of FG1 and FG2 in all inoculation conditions was detected by PCR with specific primers in 10 different samples per inoculation condition.

Figure 3.6 Venn diagram of differentially expressed genes in co-inoculated cultures. **A**, PH-1 coding sequences were used as a reference. **B**, FG1 coding sequences were used as a reference. **C**, FG2 coding sequences were used as a reference. ‘n’ represents the total number of differentially expressed genes in each condition when comparing observed and expected datasets. All genes are differentially expressed by DESeq, either up or down, by two-fold or higher, p-value < 0.05.

Figure 3.7 Genes were determined to be differentially expressed in response to isolate interaction by both RNA-Seq and qRT-PCR. *FGSG_04587* was found to be down-regulated, while *FGSG_04760*, *FGSG_09051*, and *FGSG_013870* were up-regulated as measured by RNA-Seq (white) and qRT-PCR (grey). Data was compared to expected values based on sole inoculated cultures. Asterisks denote a significance of p-value < 0.05 by DESeq.

Figure 3.8 Trichothecene genes show a trend towards down-regulation in response to isolate interaction. **A**, Analyses of *Tri* gene expression by RNA-Seq using PH-1 coding sequences as a reference; data was normalized by 'DESeq' and compared to expected values. The genes encoding for GAPDH (Gap) and β -tubulin (β -tub), which were included as control genes, are highly expressed but unaffected by interference between isolates. **B**, Analyses of *Tri6* and *Tri5* gene expression by common gene sequences and *Tri5* by isolate specific gene expression by qRT-PCR; data was compared to expected values based on sole inoculated cultures.

Table 3.1 Genes Determined to be Differentially Expressed by Two-fold or Greater, p-value < 0.05

Experimental Condition	Reference Genome		
	PH-1	FG1	FG2
FG1 vs FG2	1767	1684	1639
FG1:FG2 ; 10:1 ^a	79	67	76
FG1:FG2 ; 1:1 ^a	68	56	58
FG1:FG2 ; 1:10 ^a	61	64	55

^a experimental values were compared to expected values generated from the sole inoculum conditions

Table 3.2 Trichothecene Associated Genes have Increased Expression in FG2 Compared to FG1 by RNA-Seq

Gene Name	PH-1 as Reference		FG1 Isolate as Reference		FG2 Isolate as Reference	
	Accession in PH-1	Expression in FG2 compared to FG1 (log ₂)	Accession in FG1	Expression in FG2 compared to FG1 (log ₂)	Accession in FG2	Expression in FG2 compared to FG1 (log ₂)
<i>Tri1</i>	<i>FGSG_00071</i>	6.3 ^a	<i>g9888</i>	6.3 ^a	<i>g79</i>	6.3 ^a
<i>Tri8</i>	<i>FGSG_03532</i>	6.1 ^a	<i>g5275</i>	5.2 ^a	<i>g5383</i>	6.1 ^a
<i>Tri3</i>	<i>FGSG_03534</i>	5.3 ^a	<i>g5276</i>	5.2 ^a	<i>g5384</i>	5.3 ^a
<i>Tri4</i>	<i>FGSG_03535</i>	6.2 ^a	<i>g5277</i>	6.2 ^a	<i>g5385</i>	6.2 ^a
<i>Tri6</i>	<i>FGSG_16251</i>	4.1 ^a	<i>g5278</i>	4.1 ^a	<i>g5386</i>	4.1 ^a
<i>Tri5</i>	<i>FGSG_03537</i>	6.1 ^a	<i>g5279</i>	6.1 ^a	<i>g5387</i>	6.1 ^a
<i>Tri10</i>	<i>FGSG_03538</i>	0.2	<i>g5280</i>	0.2	<i>g5388</i>	0.2
<i>Tri9</i>	<i>FGSG_03539</i>	9.0 ^a	<i>NP^b</i>	-	<i>NP^b</i>	-
<i>Tri11</i>	<i>FGSG_03540</i>	5.1 ^a	<i>g5281</i>	5.1 ^a	<i>g5389^c</i>	5.2 ^a
<i>Tri12</i>	<i>FGSG_03541</i>	5.5 ^a	<i>g5282</i>	5.3 ^a	<i>g5389^c</i>	5.2 ^a
<i>Tri14</i>	<i>FGSG_03543</i>	6.2 ^a	<i>g5283</i>	6.2 ^a	<i>g5391</i>	6.2 ^a
<i>Tri101</i>	<i>FGSG_07896</i>	5.0 ^a	<i>g10380</i>	5.0 ^a	<i>g10702</i>	5.0 ^a
<i>GAPDH</i>	<i>FGSG_16627</i>	0.2	<i>g7821</i>	0.2	<i>g8232</i>	0.2
<i>B-Tub</i>	<i>FGSG_09530</i>	0.1	<i>g10789</i>	0.1	<i>g11116</i>	0.1

^a statistically significant by DESeq p-value < 0.05

^b 'NP' denotes genes that were not predicted by AUGUSTUS

^c the coding region of *Tri11* and *Tri12* spanned a single predicted gene in the FG2 isolate, *g5389*

Chapter 4: Distribution of a Variable Region from the *Fusarium graminearum* Genome

4.1 Abstract

Through comparative genomics, we identified genes and gene clusters contributing to the accessory genome of FHB pathogens from the *F. graminearum* species complex. In Chapters 2 and 3, we showed that clusters of genes can be present or absent in a genome, even within a single species such as *F. graminearum* (Accompanying Materials: Table S2.2). This included a set of genes on contig_207 of the 3-ADON producing isolate DAOM241165 (FG1), some of which are absent in the 15-ADON producing isolate DAOM233423 (FG2). We performed a broader investigation of genes from contig_207 and their distribution across the genus and species. Our analyses showed that genes within this contig contribute to the accessory genome for the genus and species. The genes from the contig appear to have a wide range of functions, including interactions with DNA and/or RNA, as well as proteolysis. We hypothesize that one protease *g3557* is a paralogue and resulted from a duplication of the gene *FGSG_03432*. Since these genes are expressed during infection of wheat, these genes may contribute to the infection process. Our findings substantiate the paradigm that the genomes of pathogenic fungi are diverse, and can vary in their gene content even within the same species.

4.2 Introduction

FHB is caused by a diverse set of fungi from the genus *Fusarium*, with the dominant species in North America and several other parts of the world being *F. graminearum sensu stricto* (van der Lee *et al.* 2015). It is known that within the *F. graminearum* species, isolates have phenotypic differences in disease and toxin production, as indicated in Chapter 2 (Table 2.1). In other plant pathogenic fungi, accessory genes are reported to be associated with phenotypic differences in disease, host specialization, and the production of toxic secondary metabolites (Guo *et al.* 2014; Ma *et al.* 2014). Within *F. graminearum*, phenotypic variations in toxin production have been attributed to differences in gene content or allelic composition. For example, functional copies of the genes *Tri13* and *Tri7* are required for the production of NIV and modifications thereof, and allelic differences in the gene *Tri8* contributes to differential acetylation patterns of the DON toxin, resulting in either 3-ADON or 15-ADON production (Alexander *et al.* 2011; Kimura *et al.* 2007; Lee *et al.* 2002). Data also indicate that 3-ADON producing *F. graminearum* isolates belong to a distinct genetic population that may be more competitive and are displacing a separate population of the fungus that produce 15-ADON (Kelly *et al.* 2015; Ward *et al.* 2008). In Chapter 3, we demonstrated that a 3-ADON producing isolate, DAOM241165 (FG1), is a more aggressive pathogen on wheat than a 15-ADON producing isolate, DAOM233423 (FG2), and was able to outcompete FG2 in co-inoculation experiments (Figures 3.4 and 3.5).

Comparative genomics of *F. graminearum* from Chapters 2 and 3 indicated that there are accessory genes that are absent in some individuals of the species. Moreover, the distribution of these genes across the genus and species, as well as their biological functions remains largely unknown. When investigating the accessory genome of the *F. graminearum* species complex, we

discovered that contig_207 from FG1 contained a set of 20 genes, some of which are absent in *F. graminearum* isolate FG2. We sought to investigate the gene content of this contig in detail; in particular, its distribution across the genus and species.

4.3 Results

4.3.1 Validation of the structure of contig_207 by PCR

Contig_207 was constructed by assembling Illumina sequencing reads and was predicted to contain 20 protein coding genes (*g3545-g3564*) as described in Chapter 2 (Figure 4.1 A). Assembly of short read data can be difficult, particularly in regions of the genome that are repetitive. An investigation of the structure of the contig indicated that it was approximately 90 kb in length, with about 6kb or 6.7% of the sequence consisting of ambiguous nucleotides or ‘N’s’; this is nearly ten-fold higher than the percentage of ambiguous nucleotides in the genome. The ambiguous nucleotides were located as 30 clusters that ranged from 2 to 1173 bp in length. In addition, we detected 15 simple repeat sequences by RepeatMasker that had a total length of 658 bp. The repeats ‘CCTTTAA’ and ‘CTTTAA’ spanned a 107 bp region between *g3554* and *g3555*; this sequence may be fragments from the predicted telomeric region ‘TAACCC’ for *F. graminearum* (King *et al.* 2015). The association with the telomere was further supported by a reciprocal BLASTn hit of the gene *g3554* to the gene *FGSG_14031* in the telomeric ends of chromosome 2 in PH-1 (Wong *et al.* 2011, King *et al.* 2015). Alignment of the coding sequences of *g3554* and *FGSG_14031* by MUSCLE showed 97.7% identity. However, the gene lengths of *g3554* from FG1 and *FGSG_14031* from PH-1 were slightly different, being 4260 bp and 3750 bp, respectively. The differences in length could be associated with the presences of gaps of ambiguous nucleotide sequences in both gene sequences. In the PhytoPath database assembly,

the *FGSG_14031* is flanked by a highly repetitive region rich in ‘A’ and ‘T’ nucleotides, which is indicative of telomeres (King *et al.* 2015).

Given the large number of ambiguous nucleotides and repetitive elements within the contig, we chose to confirm the structure of the contig by PCR using 35 sets of overlapping primer pairs (Figure 4.1 A). The fragments were successfully amplified using genomic DNA from FG1, indicating that the overall structure of the contig obtained by genome assembly was valid (Figure 4.1 C).

4.3.2 Distribution and evolutionary history of genes from contig_207

In order to assess the presence of genes from contig_207 in other genomes, we compared the entire gene list from FG1 to the nine other genomes from Chapter 2 by BLASTn. The results indicated that 17 out of the 20 genes were present in at least one of the other genomes, and 3 genes (*g3549*, *g3560*, and *g3561*) were present only in FG1 (Figure 4.2). The isolate NRRL28336 had the greatest number of present genes (15) from contig_207 (Figure 4.2). Since NRRL28336 and DAOM241165 are both 3-ADON producers, genes from contig_207 may be more pervasive in the producers of the 3-ADON chemotype (Figure 4.2) (Kelly *et al.* 2015). We tested this hypothesis by reference-based assembly, where the genome sequencing reads from FG2 (a 15-ADON producer) were mapped directly to the genome of FG1 (Figure 4.3). We showed that with the exception of regions within *g3554* and *g3557*, most genes had highly reduced read coverage in FG2 (Figure 4.3). This is also in agreement with BLASTn analyses to the genomes that indicate that the genes *g3554* and *g3557* were highly conserved in both FG2 and PH-1 with an expected value of 0. Analysis of *g3557* suggested that it encoded for a protease with a reciprocal BLASTn hit to *FGSG_03432*. A separate gene, *g5175*, which is present on a

different contig (contig_117), also had a strong hit to *FGSG_03432*. Upon closer inspection, we discovered that *g5175* is in a region highly syntenic to the region containing *FGSG_03432* in PH-1 on chromosome 2. This would indicate that *FGSG_03432* and *g3557* are paralogues. By quantitative PCR analysis, we were able to confirm that FG1 does possess two copies of the gene compared to FG2 (Figure 4.4). Alignment of *g3557*, *g5175*, and *FGSG_03432* by MUCSCLE showed 92.1-97.9% sequence identity between the coding sequences, with the greatest identity between *g5175* and *FGSG_03432*. We also observed variations in gene length and *g3557* was predicted to have additional sequence at the 3' end of the coding region. The predicted coding sequences of *g3557*, *g5175*, and *FGSG_03432* were 2058bp, 1125bp, and 1272bp, respectively. Together, our analyses indicate that genes from contig_207 are part of the accessory genome and that genes may originate from repetitive regions such as telomeres, or arose as a result of gene duplication.

To further understand the presence or absence of the genes from contig_207 across a greater number of fungi, we queried the genes to 30 genomes from *Fusarium* spp. by BLASTn (Figure 4.5). The genomes originated from the *F. solani* (n=2) species, *F. avenaceum* (n=3) species, *F. fujikoro* (n=9) species complex and *F. oxysporum* (n=16) species complex. BLASTn analysis showed that across 30 *Fusarium* spp. 14 out of 20 genes were detected in at least one of the 30 genomes, while we detected 8 of the genes in *F. solani*, 9 of the genes in *F. avenaceum*, 12 genes in *F. fujikoro*, and 13 genes in *F. oxysporum*. This indicated that these genes are present across the genus and may have an evolutionary history in other species. Similar analyses showed that 6 out of 20 genes were not found in any of the genomes outside of *F. graminearum*, which suggested that they may have been lost in other *Fusarium* spp. or were recently acquired in *F. graminearum*.

Some genes, such as the genes spanning *g3545* to *g3548*, were detected more often in other fusaria than in *F. graminearum*, particularly in *F. oxysporum* (present in 87.5% of genomes or greater) (Figure 4.5). This may be indicative of an early origin of these genes in the genus. The gene *g3548* was present in the genomes of only 7.3% of *F. graminearum* isolates; curiously, this gene was present in 60% of the genomes from the other *Fusarium* spp. investigated (Figure 4.5). The most common gene by both the BLASTn and PCR analyses was *g3557*. This gene was detected in all *F. graminearum* isolates and in the genomes of all distant fusaria, except in the two *F. solani* genomes.

In addition, we monitored the presence or absence of genes from *contig_207* in 54 *F. graminearum* isolates from Canada. We were also interested to understand the distribution of genes from *contig_207* in the 3-ADON and 15-ADON populations. The *Tri8* gene was used to separate the 54 isolates into the 3-ADON and 15-ADON chemotypes/populations (personal communication, Dr. Linda Harris, ORDC, AAFC). We identified 27 isolates that produce 3-ADON and 27 isolates that produce 15-ADON. Analysis showed that in addition to FG1 (a 3-ADON producer), there were two other 3-ADON producing isolates (DAOM243239, from Saskatchewan, and DAOM250000, from Manitoba) that had all 20 genes on *contig_207*. The results also indicated that most genes are more common in 3-ADON producing strains than in 15-ADON producing strains. In fact, three genes (*g3548*, *g3549*, and *g3552*) were only detected in the 3-ADON producers (Figure 4.5).

4.3.3 Functional analysis of genes from contig_207

Given their associations with the accessory genome, we were interested to know the biological function of genes in contig_207. A bioinformatics analysis by Blast-2-Go, InterPro, and pFam indicated that these genes may have roles in DNA/RNA interactions, proteolysis, membrane transport, and protein-protein interaction (Table 4.1). Of the 20 genes, genes *g3547*, *g3548*, *g3554*, and *g3563* are predicted to have helicase functions, *g3562* encodes for an endonuclease, *g3552* encodes for a protein with a bromo-adjacent homology (BAH) domain, which is associated with DNA methylation and gene regulation, and *g3553* is predicted to encode a protein with Nedd4-BP1 YacP nucleases (NYN) and Oskar and vertebrate TDRD5/TDRD7 (OST-HTH/LOTUS) domains, which are associated with RNA interactions (Table 4.1). Two genes, *g3557* and *g3560* were predicted to have protease functions, while *g3561* was predicted to have transmembrane domains and may be involved in membrane transport as part of the major facilitator superfamily of transporters. Four genes (*g3545*, *g3550*, *g3555*, and *g3559*) were predicted to encode proteins with domains such as ankyrin repeat, F-box, leucine-rich repeat, and WD40 domains, which are involved in protein-protein interactions (Table 4.1). These proteins may act as receptors or participate in larger protein complexes. The other five genes (*g3549*, *g3551*, *g3556*, *g3558*, and *g3564*) had no predicted protein domains or biological roles assigned (Table 4.1).

In addition to using predictive software, we also investigated the expression of genes from contig_207 in infected wheat tissues. The susceptible wheat cultivar ‘Roblin’ was point inoculated with spores from either FG1 or FG2, and infected spikelets were harvested 96h after infection, and the expression of genes was monitored by digital PCR (Figure 4.6). The 96h time point represented an early stage of the infection process, where the fungal mycelia have

progressed into the plant tissues and have initiated subcuticular and intercellular growth, but are typically limited to the spikelet into which they were inoculated and have not yet bypassed the node to the rachis (Boenisch and Schäfer 2011; Brown *et al.* 2010). Digital PCR is normally used to determine the absolute number of copies of a gene or cDNA in a sample by comparing the number of positive and negative amplifications of a target in emulsions PCR reaction containing approximately 20000 independent PCR reactions (Hindson *et al.* 2013). However, given that our *Fusarium* samples were extracted from wheat infected tissues, we needed to correct the amount of fungal tissues in the sample using internal controls. Our results are therefore presented as normalized counts for gene expression, similar to qRT-PCR analysis.

Results showed that after 96h of infection by FG1, expression of nine genes (*g3545*, *g3550*, *g3553-g3555*, *g3557*, *g3559-g3561*) have a normalized expression value greater than 1 in at least one replicate compared to the mock inoculated (no *Fusarium*) and no template controls, which had normalized values that were near 0 (Figure 4.6). Genes of interest or genes that demonstrated low expression (*g3546-g3549*, *g3551*, *g3552*, *g3556*, *g3558*, *g3562-g3564*) were tested using two independent sets of primers; all other genes were tested using a single primer set. Although transcript levels were low in abundance, except for *g3554*, all nine genes showed greater expression when inoculated with FG1 than with FG2, except for *g3554* (Figure 4.6). High levels of expression were detected for two genes, *g3557* and *g3554*, in wheat inoculated with FG2 (Figure 4.6). The normalized number of cDNA copies for *g3554* in FG2 was 52.2, and was higher than any of the other genes investigated, which all had normalized values below 6 (Figure 4.6). The differences of gene expression between FG1 and FG2 indicated that although the gene is present in both isolates, genes may exhibit distinct roles in the two isolates during the

infection process. Overall, the expression of these nine genes *in planta* indicated that they may have roles during the infection process.

4.4 Discussion and Conclusions

Whole genome comparisons of 10 FHB pathogens from the *F. graminearum* species complex in Chapter 2 and the more detailed direct genomic comparisons of FG1 and FG2 in Chapter 3 indicated that contig_207 is part of the accessory genome for these pathogens. These largescale comparisons indicated that this contig may be a suitable candidate for a more targeted study to ascertain its role in the accessory genome.

We successfully confirmed the structure of contig_207 from FG1, and its association with the accessory genome of *F. graminearum*. We also observed that there were a large number of ambiguous nucleotides in contig_207 compared to the rest of the genome. It is possible that at least some of the ambiguous nucleotides in the contig are due to inability of the short read assembler to resolve the repetitive regions accurately, including the potential telomeric regions adjacent to *g3554*. Nevertheless, the assembler was able to join regions flanking the gaps based on the distance between the paired-end and mate-pair sequencing reads, which outlines the potential benefits of using mate-paired libraries for the *de novo* assembly of short read data. We also provided evidence of gene duplication of one of the genes present in contig_207 in FG1. We show that *g3557* is present on contig_207 and shares identity to *FGSG_03432* that is present on chromosome 2. Cumulatively, our analyses indicate that contig_207 is part of a dynamic region of the FG1 genome and that genes within the contig may originate from multiple sources. The location of the contig in the genome of FG1 remains ambiguous; however, given the presence of telomeric-like repeats and significant diversity of the genes within the contig amongst the

genomes of FHB pathogens, it is possible that this contig is located in a repetitive region of high nucleotide variability such as telomeres and nucleotide variation hotspots within chromosomes that we identified in Chapter 2 and that have been shown in other studies (Figure 2.3) (Cuomo *et al.* 2007; Gardiner *et al.* 2012).

The presence and absence of genes between *Fusarium* species and within the *F. graminearum* populations that produce either 3-ADON or 15-ADON, give insights into the evolutionary history of the contig as well as specific genes within the contig. The observation that there were genes that were more common within *Fusarium* species other than *F. graminearum* may indicate a more distant evolutionary origin; four of these genes (*g3545-g3548*) were also co-localized at the end of the contig and may have been lost or acquired together in their history. The gain and loss of clusters of genes has been reported previously for *Fusarium*, specifically in the context of secondary metabolism and disease (Gardiner *et al.* 2012; Proctor *et al.* 2009; Proctor *et al.* 2013). There were three other genes (*g3553, g3559, g3562*) that were also more common in the genus *Fusarium* than within specific *F. graminearum* chemotypes/populations; however, these genes were more broadly distributed across the contig. Together, this may indicate that there is a long history of evolution of these genes from contig_207 and that the genes may have been acquired and/or lost over time through many independent events. Within the species, genes from contig_207 were more commonly associated with 3-ADON producing isolates, with three genes not detected in any of the 15-ADON producing strains. It is possible that these three genes underwent a gain or loss near or after the separation of these two distinct genetic populations (Kelly *et al.* 2015). It is also possible that some of the genes contain increased sequence variation that limited their detection by PCR. Directed sequencing of the genes in other isolates, or a comparison of these genes across

additional genomes as their entire genome sequences become available, would clarify the specific events that led to the current structure of contig_207 in FG1.

Although not all genes contain known functional domains, their expression pattern during infection suggests a potential role in this process. Of particular interest is that both *g3557* and *g3560* are expressed during infection and encode proteases; proteases have been shown to degrade proteins involved in the initiation of defense response pathways in plant hosts (Jones and Dangl 2006). As we documented, *g3557* has been duplicated and was present as *FGSG_03432* on chromosome 2 and it would be interesting to know which of these gene copies are in fact contributing to the transcripts detected during infection and if the gene in contig_207 is expressed differently as a result of its location in the genome. Curiously, none of the five uncharacterized genes had high levels of gene expression *in planta*, indicating that either they are not expressed or they are miscalled genes by the predictive software, though *g3556* and *g3564* were shown to have low levels of expression in *in vitro* by RNA-Seq analyses in Chapter 3 (Accompanying Materials: Table S3.3). Functional analyses of the genes from the contig_207 would elucidate their functions, specifically, if they contribute to the aggressiveness of the emerging isolates from the 3-ADON chemotype (Ward *et al.* 2008; Kelly *et al.* 2015).

In summary, we have explored a subset of genes from the accessory genome of *F. graminearum* and its distribution across the genus and additional members of the species. Our results indicate that accessory genes can be highly variable and their presence or absence are likely determined through a long history of acquisition/loss, rearrangement, and duplication. In addition, some of these genes are expressed during infection of wheat and could therefore have important biological roles during host colonization. From a broader evolutionary perspective, our results indicate that despite the genomic profile of *F. graminearum* that suggests a low level of

genome plasticity (Ma *et al.* 2010), these important plant pathogens do have the potential to alter their gene content. The genes from contig_207 and other genes from the accessory genome require additional attention to delineate potential roles in evolution and adaptation in these important plant pathogens.

4.5 Materials and Methods

4.5.1 Analyses of contig structure and distribution

The sequence for contig_207 from DAOM241165 (FG1) was analyzed for repetitive elements by RepeatMasker by the available web server (<http://www.repeatmasker.org/>) using the cross_match search engine with slow speed/sensitivity settings. The simple repeat sequences that were identified were then compared to known repeat sequences. PCR was used to confirm the structure of the contig using 35 sets of overlapping primers pairs (Accompanying Materials Table S3.4).

To validate the lack of conservation of contig_207 in DAOM233423 (FG2), a reference based assembly was performed using CLC Genomics Workbench. BLASTn was used to further investigate the presence and absence of specific genes across the genomes of other fusaria; genes were considered present in a given genome if the BLASTn expected value was $<1E-10$. The BLASTn results were used to extrapolate regions within the genes that had increased conservation across multiple genomes, which was used to design primer pairs for each gene. PCR was then used to screen these primer pairs across 54 genomic DNA samples from *F. graminearum* isolates obtained across Canada. The copy number of the protease gene *FGSG_03432* and its paralogue *g3557* were determined by quantitative PCR using the *Tri6* gene

as an endogenous control (Walkowiak and Subramaniam 2014). A complete list of primer sequences is provided (Accompanying Materials S3.4).

4.5.2 Functional analyses of genes from contig_207

To further investigate the functions of genes from contig_207 we performed *in silico* analyses of gene functions by Blast-2-Go, as described in Chapter 2. We also performed *in planta* analyses of the genes by determining their expression during infection of wheat. Expression was determined by point inoculated spores from FG1 or FG2 into the susceptible wheat cultivar ‘Roblin’ as described in Chapter 3. After 96 h, the infected wheat spikelets were harvested and flash frozen in N₂ (liquid). Three biological replicates were processed for both FG1 and FG2; each replicate consisted of five infected spikelets from separate wheat heads. Each replicate was subject to RNA isolation and cDNA synthesis as described in Chapter 3. Expression of the 20 genes was determined by Bio-Rad qX200 Droplet Digital PCR according to manufacturer instructions. The genes for β -tubulin, glyceraldehyde 3-phosphate dehydrogenase (GAPDH), and elongation factor 1 (EF1 α) were used as internal controls to quantify the amount of *Fusarium* cDNA in the samples and was used to normalize the data for relative comparison between samples. A complete list of primer sequences is provided (Accompanying Materials S3.4).

Figure 4.1 Confirmation of the structure of contig_207 from FG1. **A**, The position of 20 genes along contig_207 as predicted by AUGUSTUS. **B**, The approximate positions of 35 overlapping primer pairs used to confirm the structure of contig_207. **C**, PCR and gel electrophoresis confirmation of the successful amplification of the 35 overlapping primer pairs.

Figure 4.2 Distribution of genes from contig_207 across the genomes of nine other closely related genomes from other FHB pathogens. Genes from contig_207 that are present (black) or absent (white) in 9 other genomes were determined by BLASTn with an expected value $<1E-10$ for presence.

Figure 4.3 A reference based mapping of sequencing reads from FG2 directly to contig_207 of FG1. Reads that could be mapped to contig_207 and their coverage are presented.

Figure 4.4 Relative copy number of the protease gene *FGSG_03432* or *g3557*. Quantitative PCR was used to determine the relative number of copies of the protease gene *FGSG_03432* or *g3557* and *Tri6* was used as an endogenous control with a copy number of one.

Figure 4.5 Percent overlap of genes from contig_207 in other *Fusarium* genomes. Thirty genomes of *Fusarium* spp. distantly related to *F. graminearum* were assessed for the presence of genes from contig_207 by BLASTn and were considered present if the gene had an expected value $<1E-10$ (grey bars); the genomes originated from the *F. solani* (n=2), *F. avenaceum* (n=3), *F. fujikoroii* (n=9) and *F. oxysporum* (n=16) species or species complexes. For *F. graminearum*, the presence of genes were determined by PCR using gene specific primers for both 3-ADON producing strains (n=27) (black bars) and 15-ADON producing strains (n=27) (white bars).

Figure 4.6 Normalized cDNA copies for genes from contig_207 after 96 h of infection in wheat. Digital PCR was used to determine the expression of each gene for both FG1 (black bars) and FG2 (white bars). Three internal control primer sets corresponding to the genes EF1 α , β -tubulin and GAPDH were used to normalize for the amount of fungal cDNA within samples. Error bars reflect the maximum value determined based on three biological replicates. Genes marked with an asterisk (*) have a normalized number of cDNA copies greater than 1 in at least one replicate.

Table 4.1 Predicted Biological Roles for Genes from Contig_207

Gene	Protein Motifs ^a	Gene Ontology IDs ^b
DNA/RNA binding (replication, transcription and gene regulation, epigenetics, maintenance)		
<i>g3546</i>	-	F:DNA binding
<i>g3547</i>	DEAD/DEAH ^c box helicase	F:ATP binding; F:nucleic acid binding
<i>g3548</i>	-	P:biological_process; F:helicase activity; F:nucleic acid binding; F:ATP binding; F:ATP-dependent helicase activity
<i>g3552</i>	BAH	F:chromatin binding
<i>g3553</i>	NYN, OST-HTH/LOTUS	-
<i>g3554</i>	Sucrose non-fermenting 2 (SNF2), helicase	F:DNA binding; F:ATP binding
<i>g3562</i>	endonuclease	F:zinc ion binding; F:DNA binding
<i>g3563</i>	-	F:nucleic acid binding; F:ATP binding; F:ATP-dependent helicase activity; F:DNA binding
Proteolysis		
<i>g3557</i>	aspartic peptidase	F:aspartic-type endopeptidase activity; P:proteolysis
<i>g3560</i>	peptidase s8	F:serine-type endopeptidase activity; P:proteolysis
Membrane transport		
<i>g3561</i>	-	C:integral to membrane; P:transmembrane transport
Protein-protein interaction (protein complex formation, receptors)		
<i>g3545</i>	ankyrin repeats	F:protein binding; F:catalytic activity; P:nucleoside metabolic process
<i>g3550</i>	ankyrin repeats	F:protein binding; P:nucleoside metabolic process; F:catalytic activity
<i>g3555</i>	F-box, leucine-rich repeat	F:protein binding
<i>g3559</i>	NAIP CIITA HET-E and TP1 (NACHT), WD40 ^c	F:protein binding
Unknown function		
<i>g3549</i>	-	-
<i>g3551</i>	-	-
<i>g3556</i>	-	-
<i>g3558</i>	-	-
<i>g3564</i>	-	-

^a All protein domains were predicted by pFam or by Interpro (unintegrated domains were excluded)

^b Gene Ontologies predicted by Blast-2-Go are organized into molecular functions (F), biological processes (P), and cellular components (C)

^c Motif name is based on the standard single letter amino acid code

Chapter 5: Genomics of *Fusarium avenaceum*

5.1 Abstract

The filamentous ascomycete fungus *Fusarium avenaceum* is a pathogen of several plant species and can cause the Fusarium Head Blight (FHB) disease of wheat in Canada. This fungus is part of the *F. tricinctum* species complex and is distantly related to the more common FHB pathogens from the *F. graminearum* species complex. We report on the genome sequences of three *F. avenaceum* isolates, two from Canadian wheat, and one from barley in Finland. Our analyses indicate that the genome can vary in size from 41.6 to 43.1 Mb and is composed of eleven supercontigs that may correspond to entire chromosomes. There are also 13217 to 13445 predicted protein-coding genes, 95% of which were common between genomes.

5.2 Introduction

Currently, the prominent FHB pathogen that infects Canadian wheat crops is *F. graminearum*; however, infection of wheat by *F. avenaceum* is common and was likely even more prevalent prior to the 1990's (Clear and Patrick 2000; Clear and Patrick 2010; Graefenham *et al.* 2014). It has been suggested that as part of the dynamic changes in pathogen profiles over the years, *F. graminearum* has emerged as the dominant FHB pathogen and is more aggressive than other FHB pathogens, such as *F. avenaceum* (Clear and Patrick 2000). In contrast to many of the other FHB pathogens, such as *F. graminearum*, *F. culmorum*, *F. sporotrichioides*, or *F. poae*, which are closely related, *F. avenaceum* is part of a more distant taxonomic group, the *F. tricinctum* species complex.

Similar to *F. graminearum*, *F. avenaceum* is globally distributed and can produce a wide array of secondary metabolites that contaminate infected crops, including beauvericin and

enniitins (Desjardins 2006; Logrieco *et al.* 2002). In contrast to *F. graminearum*, which is suggested to have a fairly narrow host range and is mostly observed on cereals, *F. avenaceum* has been reported to have a very broad host range and can cause disease on a number of different hosts, including 160 different plant genera from 26 families (Desjardins 2006). There can also be significant differences in host aggressiveness between individual isolates of *F. avenaceum* in different hosts (Feng *et al.* 2010; Nalim *et al.* 2009). *F. avenaceum* also does not appear to have the gene *Tri5*, trichodiene synthase, which is required for trichothecene biosynthesis (Desjardins 2006; Leslie and Summerell 2006). This is curious considering trichothecene production has been demonstrated to be essential for FHB infection by *F. graminearum* (Jansen *et al.* 2005). The biology of *F. avenaceum* infection is not well described and there are several fundamental questions that remain to be answered, such as: Why is *F. avenaceum* less aggressive than *F. graminearum*? What are the important factors contributing to FHB infection by *F. avenaceum* and are they similar or different than those used by *F. graminearum*? What are the mechanisms contributing to the broad host range by this pathogen?

Much of the research that has gone into understanding *F. graminearum* biology has been facilitated by the available genome sequence for isolate PH-1. The goal of this study was to provide a genome sequence for the future genetic dissection of *F. avenaceum*. This FHB pathogen has clear differences in disease and secondary metabolism properties compared to *F. graminearum*. The availability of its genome sequence provides insights into possible infection strategies of this pathogen, as well as its relations to FHB pathogens from the *graminearum* species complex.

5.3 Results

5.3.1 Genome structure and alignments of *F. avenaceum* genomes

The *de novo* assembly of three *F. avenaceum* isolates resulted in 77-104 scaffolded contigs with N50 values ranging from 1.43Mb to 4.14Mb, with some contigs assembling as large as 6.5Mb (Table 5.1). The two Canadian genomes (FAV1 and FAV2) had the greatest N50 values, likely a consequence of the higher sequencing depth; the genomes had coverage of 426.6 and 986.2 fold compared to the 21.6 fold coverage of the genome of the Finnish isolate (FAV3). The overall length of the three genomes was very similar, ranging from 41.6 to 43.2 Mb, which is larger than the genomes of the FHB pathogens from Chapter 2 by roughly 5 Mb.

Unlike the genomes from Chapter 2, which were highly related and could be ordered according to the reference sequence of isolate PH-1; these newly sequenced genomes required additional analyses to obtain a higher order assembly, which was facilitated by synteny between genome alignments. A comparison of the genomes revealed large syntenic regions with 91–96% nucleotide alignment and few unaligned bases, inversions, or translocations, particularly between the two Canadian isolates (Figure 5.1A). The synteny between contigs was used to order contigs between genomes. Reiterative reordering of the contigs between the two Canadian genomes resulted in a stable arrangement of the contigs and alignment overlaps. This facilitated the construction of eleven supercontigs that ranged in size from approximately 0.8 Mb to 6.5 Mb (Figure 5.2). The size and quantity of supercontigs were comparable to karyotyping results that have been previously described for this species (Lysøe *et al.* 2014).

Direct comparison of *F. avenaceum* genomes reveal genomic diversity among isolates within the species. When comparing the three *F. avenaceum* genomes to each other, the genomes

were highly syntenic; however, a potential translocation event was observed between scaffolded contig_3 from the Canadian isolate FAV1 and contig_2 from the Finnish isolate FAV3 (dotted box Figure 5.1C). The translocated region was almost entirely within a single contig in the second Canadian isolate FAV2 (contig_22); as such, the contig from FAV2 can be arranged with little conflict to either of the two other genomes. In addition to possible structural differences between the genomes, approximately 1.4% and 3.2% of the aligned nucleotides exhibited single nucleotide polymorphisms (SNPs), insertions, or deletions between the two Canadian and the Canadian and Finnish isolates, respectively. These genetic differences were unevenly distributed across the genomes and were largely concentrated at the ends of the supercontigs (Lysøe *et al.* 2014). The co-localization of genetic variations is similar to what has been reported previously in the alignment of other FHB pathogens from the *F. graminearum* species complex in Chapter 2 (Figure 2.3), as well as other studies (Cuomo *et al.* 2007; Gardiner *et al.* 2012).

When we compared the supercontigs of FAV1 to the chromosomes of *F. verticillioides* and *F. graminearum*, there were large syntenic blocks that arranged linearly to regions of chromosomes, particularly for *F. verticillioides* (Figure 5.3A). For example, supercontig 1 from FAV1 appears to correspond nearly perfectly to chromosome 1 from *F. verticillioides*, and is also syntenic to a portion of chromosome 1 from *F. graminearum* after some rearrangements and inversions (Figure 5.3B). Together, the alignments support the structure of the supercontigs and indicate some level of synteny across the genomes of these species. As indicated in Chapter 2, there was at least 12% alignment between *F. avenaceum* and the genomes of *F. graminearum*, *F. asiaticum*, and *F. meridionale* when NUCMER was used; however, alignment increased to at least 51% when the less stringent aligner PROMER was used (Figure 2.1).

5.3.2 Comparisons in gene content between *F. avenaceum* genomes

In addition to comparing genome structure between *F. avenaceum* isolates, we were also interested in the gene content of these newly sequenced fungi and the level of genetic overlap between isolates. AUGUSTUS was used to predict genes in the genomes, which resulted in 13445, 13293, and 13217 predicted genes in FAV1, FAV2, and FAV3, respectively. BLASTn analysis indicated that more than 95% of predicted genes had a significant hit within the two other *F. avenaceum* genomes, with greater overlap existing between the two Canadian genomes than the Canadian and Finnish genomes (Figure 5.4). Genes that differ between the genomes were annotated to have functions in secondary metabolism, including polyketide synthase and nonribosomal peptide synthase genes; differences in the production of secondary metabolites were also reported between the isolates (Lysøe *et al.* 2014). Other genes that differed between genomes include mating-type idiomorphs, which indicate that the species has two clear mating-types and likely exhibits a heterothallic or outcrossing sexual lifestyle. Together, the *F. avenaceum* isolates have clear differences in genome structure and gene content that may be associated with isolate specific adaptations. The level of genetic overlap is similar to what was observed between the genomes of the other closely related FHB pathogens (Figure 2.6).

We performed comparative BLASTn analyses to better understand the genetic overlap between the three *F. avenaceum* isolates and ten isolates from the *F. graminearum* species complex from Chapter 2. The results indicated that 80.6-81.3% of the genes from the genome of a given *F. avenaceum* genome had a successful BLASTn hit in the genome of at least one of the genomes from the *F. graminearum* species complex. At the same time, when the pan genome of the ten isolates from the *F. graminearum* species complex was compared to the genomes of the three *F. avenaceum* isolates, 74.7-75.1% of genes had a successful BLASTn hit. A separate

analysis by others that compared FAV3 to *F. graminearum* PH-1 genes (MIPS v3.2) by BLASTp determined that 10140 or 76.7% of genes had reciprocal hits, indicating that there are over 3000 genes in each species which may have diverged or were acquired/lost (Lysøe *et al.* 2014). *F. avenaceum* genes that are absent in *F. graminearum* were determined by gene ontology to be functionally enriched in redox processes, metabolism (i.e. toxin and cell wall metabolism), and transcription from RNA polymerase II promotor (Lysøe *et al.* 2014).

5.4 Discussion and Conclusions

The genome sequencing of three *F. avenaceum* isolates from the *F. tricinctum* species complex provide an excellent framework for the understanding of these FHB pathogens that belong to a distinctly different taxonomic group from the FHB pathogens that are from the *F. graminearum* species complex (Chapter 2). Assembly and comparisons of three similar *F. avenaceum* genomes allowed for the construction of a higher order structure for the genome of *F. avenaceum*. Genome comparisons within the species showed trends of nucleotide variation as well as potential structural changes between the genomes, particularly between the Finnish and Canadian isolates. Additional experimentation would be required to validate the assembly structure in the proposed translocated region to determine the correct arrangement of this contig in FAV2, and if the discrepancy between FAV1 and FAV3 is truly a translocation or the result of genome misassembly. Nevertheless, the high level of genomic and genetic overlap between the isolates indicate that these isolates are closely related and that biological pathways are mostly shared between them (~95% gene overlap or greater). Although there is a broad host range for *F. avenaceum*, population genetics indicate that diversity exists within the species, and that host specificity may be independent of genetic population (Nalim *et al.* 2009).

When comparing gene content of the *F. avenaceum* isolates to FHB pathogens from the *F. graminearum* species complex, there was a high level of overlap (>75% gene overlap). This indicates that most genes are shared across their evolutionary histories. However, when considering the genomic alignments, there appears to be several largescale structural changes that must have occurred in their genomes throughout the evolution of these pathogens. In addition to the genetic and genomic differences between the fungi having potential to contribute to pathway differences involved in toxin production in disease, they also contribute to other aspects of the general biology of these fungi. For example, *F. graminearum* is homothallic and individuals can complete the sexual lifecycle because both mating type loci are present in every individual, whereas *F. avenaceum* is heterothallic and requires two individuals of opposing mating types to complete the sexual lifecycle; this is supported by the presence of two mating types, only one of which is present in a given isolate. These observations are in agreement with previous findings (Feng *et al.* 2010).

F. avenaceum has a wider host range than its counterpart *F. graminearum*. Furthermore, *F. avenaceum* does not require the trichothecene product DON to colonize wheat. The genomic analysis presented in this chapter provides a framework for the genetic dissection of *F. avenaceum* to uncover genes important for disease, as well as the production of unique secondary metabolites and toxins.

5.5 Materials and Methods

5.5.1 Genome sequencing and comparative genomics

The two Canadian *F. avenaceum* strains, FAV1 (FaLH27, DAOM242378) (winter wheat, Nova Scotia) and FAV2 (FaLH03, DAOM242076) (spring wheat, New Brunswick) were isolated from wheat samples harvested in 2011 (Canadian Grain Commission, Winnipeg, MB). Genomic DNA was extracted from fungal mycelia and was sequenced at the Genome Quebec Innovation Centre (Montreal, Quebec) by Illumina HiSeq platform (100bp paired-end). Reads were assembled in CLC Genomics Workbench and optimized for contig length as in Chapter 2. The higher order assembly was resolved by reiteratively reordering the contigs between the Canadian isolates by ABACAS until the order of the contigs stabilized. Synteny and boundaries of overlapping regions from the genome alignments by MUMMER were used to finalize higher order assembly. Gene prediction and comparisons in gene content were performed with AUGUSTUS and BLASTn as in Chapter 2.

Figure 5.1 Genome alignment of the three *F. avenaceum* isolates (FAV1, FAV2, and FAV3) reveals that the genomes are highly syntenic. **A**, Alignment of the genomes from the two Canadian isolates (FAV1 and FAV2), **B**, a Canadian (FAV2) to a Finnish (FAV3) isolate, and **C**, the separate Canadian isolate (FAV1) to the Finnish (FAV3) isolate are shown. Aligned regions are represented in red, while aligned inversions are represented in blue; unaligned regions are represented as gaps or breaks in the lines. Regions aligning away from the diagonal are translocations. A potential translocation event between the Canadian and Finnish isolates are highlighted by a dotted box.

Supercontig Size		Isolate	Contig
1	6.5Mb	FAV1	3
		FAV2	5 22 17 23
		FAV3	19 22c 34 2c 38 37c 28c 15c
2	5.7Mb	FAV1	4
		FAV2	13 9
		FAV3	26 1 24
3	4.9Mb	FAV1	8 14
		FAV2	2
		FAV3	23c 7 64c 3c
4	4.8Mb	FAV1	2
		FAV2	3
		FAV3	73c 51 33 13 18c 9 74 40c 27
5	4.6Mb	FAV1	1
		FAV2	4
		FAV3	43 25c 4 52 82 5 54c
6	4.4Mb	FAV1	10 19
		FAV2	15 10 12
		FAV3	6c 11c 20
7	4.1Mb	FAV1	24 26 11 20 30
		FAV2	7
		FAV3	12c 30 29 31 8
8	2.8Mb	FAV1	13
		FAV2	11
		FAV3	32c 41 16 14 81c 48 63c
9	2.0Mb	FAV1	12
		FAV2	16 29 21
		FAV3	21 17
10	1.5Mb	FAV1	25 5
		FAV2	6
		FAV3	10c
11	0.8Mb	FAV1	9
		FAV2	8 18
		FAV3	39 44 45 56c 55 49 36

Figure 5.2 Supercontigs of the three *F. avenaceum* isolates; two from Canada (FAV1, FAV2) and one from Finland (FAV3). The corresponding contigs from each genome within each supercontigs are presented. The size of the supercontigs is based on the contig lengths from FAV2.

Figure 5.3 Alignment of higher order assembly from FAV1 to *F. verticillioides* and *F. graminearum* chromosomes. **A**, Alignment of the FAV1 and *F. verticillioides* genomes, and **B**, alignment of the FAV1 and *F. graminearum* (strain PH-1) genomes.

FAV3 (n=13217)

FAV1 (n=13445)

FAV2 (n=13293)

Figure 5.4 Overlap in predicted genes for the three *F. avenaceum* genomes (FAV1, FAV2, and FAV3). Genes predicted by AUGUSTUS were compared to the genome sequences of the other two isolates by BLASTn and were considered present in the other genomes when having a BLASTn expected value $<1E-10$. 'n' signifies the number of predicted genes in each of the genomes. Values that are yellow, green, and blue correspond to predicted genes from FAV1, FAV2, and FAV3, respectively.

Table 5.1 Genome Statistics for *F. avenaceum* Genomes

Strain	FAV1 (LH27)	FAV2 (LH03)	FAV3 (05001)
Origin	Canada	Canada	Finland
Sequencing Technology	HiSeq	HiSeq	454
Genome Size (Mb)	43.2	42.7	41.6
Sequencing Coverage (fold)	986.2	426.6	21.6
Number of Scaffolds	77	104	83
N50 (Mb)	4.14	4.11	1.43
GC content (%)	48%	48%	48%

Chapter 6: Final Discussion

6.1 Summary of Findings

This thesis used a combination of largescale analyses as well as more focused and directed investigation, to enhance our knowledge of historical and emerging pathogens that cause FHB in Canada and around the world. We performed whole genome sequencing and comparative genomics of FHB pathogens, including eight new genomes from the *F. graminearum* species complex, as well as two new genomes from the *F. tricinctum* species complex. We identified genes that are in common between isolates, as well as accessory genes that are present only in a subset of genomes and may be unique to a species or an isolate. Several genes that we identified appear to be dynamic in FHB pathogens encode proteins that have structural signatures associated with genes previously characterized to be involved in toxin production, including, polyketide synthetases, non-ribosomal peptide synthetases, and terpene synthetases. We also demonstrated that many of the variable genes are expressed during infection of wheat or in culture conditions used to induce the production of toxins. The 1827 accessory genes identified in Chapter 2 provide an excellent resource for future studies with respect to gene diversity and dynamics, specifically pertaining to isolate and species adaptations within the *F. graminearum* species complex.

We provided evidence that showed that DAOM241165, a 3-ADON producer, is more competitive than DAOM233423, a 15-ADON producer. We also showed that intraspecies interactions between these two *Fusarium* isolates can be disadvantageous to the fungi, resulting in reduced toxin production and disease. Transcription profiling by RNA-Seq identified candidate genes that are affected by these interactions, including genes associated with trichothecene production. Detailed analyses of the genomes of the two *Fusarium* isolates

revealed that a contig (contig_207 in DAOM241165) with 20 predicted genes is not conserved in DAOM233423 and several isolates from the *F. graminearum* species complex, with many genes considered absent. In fact, none of the genes from contig_207 could be computationally transferred to the genome of DAOM233423 by the Rapid Annotation Transfer Tool (RATT). The large number of ambiguous nucleotides, repetitive elements, and possible associations with the telomere likely contributed to the lack of synteny in the alignments that prevented genes from being transferred computationally. Nevertheless, detailed analysis of this contig, its structure, and its distribution across the species and genus, confirmed its association with the accessory genome. Further investigations of this region and its genes sets precedence to show that gene content can vary between these pathogens across a greater number of isolates.

The approaches that we used to profile the FHB pathogens have been used in other systems to identify genes associated with disease and pathogen evolution (Baltrus *et al.* 2011; Ma *et al.* 2010; Schmidt *et al.* 2016). The concept of pan, common, and accessory genomes have been successfully applied to the study of bacteria to identify genes that contribute to niche adaptation and disease (Baltrus *et al.* 2011). Studies conducted on prokaryotes have led to the development of a model where genes associated with adaptive traits are co-localized in gene islands that could be readily transferred between bacteria (Baltrus *et al.* 2011). Similar to prokaryotes, horizontal transfer has also been demonstrated in fungi with the transfer of an accessory chromosome in *F. oxysporum* that confers pathogenicity to tomato (Ma *et al.* 2010). The transfer of an accessory chromosome in *F. oxysporum*, as well as other hints of smaller scale horizontal gene transfer and gene gain or loss events in *F. graminearum* and *F. pseudograminearum* indicates that fungal genomes can change in their structure and gene content (Gardiner *et al.* 2012; Ma *et al.* 2010; Proctor *et al.* 2013; Sieber *et al.* 2014).

There are several lines of evidence that indicate that the genome of *F. oxysporum* may be more amenable to change than the genome of FHB pathogens from the *F. graminearum* species complex: (i) it is larger (60Mb vs. 36Mb), (ii) has more repetitive sequences (3.98% vs. 0.03% genome is transposable elements and 16.83MB vs 0.24Mb of the genome is repetitive elements), and (iii) it contains more chromosomes (15 vs. 4) (Guo *et al.* 2014; Ma *et al.* 2010; Ma *et al.* 2014; Schmidt *et al.* 2016). In addition, *F. oxysporum* is diverse in its niches and host specializations, possibly as a consequence of its plastic genome and the presence of dispersible chromosomes (Ma *et al.* 2010). In contrast, FHB pathogens from the *F. graminearum* species complex have a narrow host range, typically infecting cereals and their relatives, though phenotypic differences in disease and secondary metabolism have been observed for these fungi (O'Donnell *et al.* 2000; Starkey *et al.* 2007).

This study shows that despite the stable appearance of the genomes of *F. graminearum*, there is some level of plasticity, which allows for genes to be acquired, lost, or duplicated; particularly in telomeric regions and select regions within chromosomes that appear to be hot spots for genomic differences. These regions likely experience higher recombination and horizontal gene transfer rates, thereby causing increased gene turnover. Differences in these regions have been reported previously and genes within these regions have greater expression during infection of wheat, thus correlating regions of nucleotide variability with the accessory genome and pathology (Cuomo *et al.* 2007; Gardiner *et al.* 2012; Lysøe *et al.* 2011). Together, these genes are good candidates for future study for virulence and/or adaptation to specific environments or hosts.

This thesis also provides genomic resources to study *F. avenaceum*. Historical records show that *F. avenaceum* was a common cause of FHB in Canada before being displaced recently

by *F. graminearum* (Clear and Patrick 2010; Graefenham *et al.* 2014). The infection strategy of *F. avenaceum* towards wheat remains under explored. It is known that trichothecenes contribute to FHB infection by *F. graminearum*; however, there are no reports of *F. avenaceum* isolates that produce trichothecenes (Desjardins 2006; Jansen *et al.* 2005). It is possible that other effectors or alternative secondary metabolites are produced by *F. avenaceum* to facilitate the initiation of infection and the subsequent spread of the pathogen within wheat. The genomic data provide researchers with a framework to test these hypotheses and characterize genetic pathways involved in secondary metabolism and disease associated with FHB and other disease caused by these pathogens. By better characterizing *F. avenaceum* and comparing infection strategies, we can better understand why the emerging *F. graminearum* pathogens are dominating as the FHB pathogen in Canada and other regions around the globe.

To summarize, we have shed light on the genomic and genetic differences of FHB pathogens from the *F. graminearum* and *F. tricinctum* species complexes, and provided candidate genes that may contribute to their known biological diversity. We demonstrated that fungal isolates can outcompete each other during infection of wheat, and identified candidate genes that may contribute to intraspecies interactions, as well as the negative consequences that these interactions have on fungus' ability to cause disease and produce toxins.

6.2 Studying FHB Pathogens in a Modern Genomics Era

We have shown that whole genome and transcriptome analyses by next-generation sequencing can be used to profile FHB fungi and their interactions. We applied analysis tools from independent sources, including but not limited to the gene prediction software AUGUSTUS, as well as tools from software packages, such as CLC Genomics Workbench and the Post Analysis Genome Improvement Toolkit (Stanke and Waack 2003; Swain *et al.* 2012). Using these tools, we were able to identify genetic variation that could be contributing to phenotypic differences between isolates. As analytical tools for sequencing data become more accessible, it will become easier to make more informed conclusions from sequencing studies and data from multiple studies can be better integrated.

As sequencing data pertaining to FHB pathogens continues to accumulate, it will be imperative that the data be managed and referenced in appropriate databases so that the data can be accessed by the research community. As we showed, it may be more suitable to combine datasets to develop a pan genome for a given species (Chapter 2), as opposed to having a single well annotated reference genome, which is currently the case for the PH-1 genome for *F. graminearum*. An advantage of constructing a pan genome is that it reflects a total number of genes that a species has evolved. Moreover, any emergence of new species can be monitored with the addition or omission of genes in the pan genome. It will be of interest to the research community to integrate the pan genome to complement other databases, such as the KEGG biochemical pathway database, protein-protein interaction networks, catalogues of metabolic gene clusters, or gene expression databases and networks (Guo *et al.* 2016; Hansen *et al.* 2015; Lawler *et al.* 2013; Lysøe *et al.* 2011; Sieber *et al.* 2014; Zhao *et al.* 2009; Zhao *et al.* 2013). Combining these largescale systems biology datasets as part of data mining studies may be the

key to identifying genes that are associated with disease and toxin production in FHB fungi (Jia and Tang 2015; Sperschneider *et al.* 2015). With the release of the draft genome for bread wheat, similar genomics analyses of the host with respect to adaptation to emerging pathogens and climate changes can also be conducted (Mayer *et al.* 2014).

6.3 Limitations of Largescale Genomics Studies

Although the largescale systems biology datasets have been a valuable resource for understanding the biology of FHB pathogens and their hosts, these datasets have limitations. The available genome sequence for *F. graminearum* isolate PH-1 has guided countless research studies; however, a major challenge that remains is that >50% of genes remain functionally unclassified (Wong *et al.* 2011). Data mining and merging multiple largescale systems biology datasets, including datasets from other organisms, may produce some insights into these unclassified genes, but determining the true function of genes requires functional characterization of individual genes through genetic analysis. Functional characterization of genes involves gene disruption or constitutive expression. This can be complemented by additional phenotypic analyses by microscopy, biochemistry, and disease monitoring in the case of FHB pathogens. Although laborious and costly, functional validation by physical experimentation is the best way to understand the functions of genes.

Technical advances to functionally characterize genes have made the process more systematic and streamlined. For example, historically, transgenic or mutant *F. graminearum* strains were constructed by the laborious method of transforming protoplasts made from actively growing mycelia. Now, conidia can be transformed by *Agrobacterium* mediated methods with remarkable efficiency (Frandsen *et al.* 2008). It may also be possible to integrate new gene targeting techniques, such as CRISPR/Cas9 to study genes in *F. graminearum* (Jia and Tang

2015). For example, this technique could be adapted to transfer the entire *contig_207* from FG1 strain into the FG2 strain and observe for phenotypic changes.

Largescale genetic studies targeting entire classes of genes, such as transcription factors, kinases and phosphatases has allowed us to associate many of these genes with toxin production and disease (Son *et al.* 2011, Wang *et al.* 2011b; Yun *et al.* 2015). For example, the study by Son *et al.* disrupted 657 transcription factors and assessed 17 phenotypes for each mutant; 170 of the mutants investigated had observable phenotypic differences, including 55 with reduced DON production and 62 with reduced disease, though the downstream targets of these genes are not well understood (Son *et al.* 2011). Individual studies have also linked specific genes to disease and toxin production, which include genes involved in primary metabolism, secondary metabolism, as well as stress response (Gardiner *et al.* 2009; Hou *et al.* 2002; Lee *et al.* 2002; Subramaniam *et al.* 2015; Walkowiak and Subramaniam 2014; Wang *et al.* 2014). To understand the mechanism by which these genes are associated with toxin production and disease, integration of other datasets are required, such as protein-protein interaction studies, gene and protein expression analysis, as well as metabolic profiling. For example, our research team undertook a largescale chromatin immunoprecipitation (ChIp-Seq) experiment with the positive regulator of the trichothecene pathway, Tri6, as bait. This led to the identification of potential regulatory targets and the DNA binding site for Tri6 (Nasmith *et al.* 2011). Additional functional characterization of regulatory targets established a link between Tri6 and primary metabolic pathways (Subramaniam *et al.* 2015). We envisage that similar strategies could be employed to study other transcription factors identified by Son *et al.* involved in both toxin and disease regulation in *F. graminearum*.

Thus far, approximately 10% of *F. graminearum* genes have been functionally characterized in some form (Jia and Tang 2015). Many of the functionally characterized genes from *F. graminearum* have been catalogued in the PHI database (Urban *et al.* 2014). However, a more comprehensive database combining functional data and largescale systems biology data would be ideal and is currently underway using the platforms at PhytoPathDB (<http://www.phytopathdb.org/>) and FungiDB (<http://fungidb.org/>) (King *et al.* 2015). These databases are designed to store and analyze large biological datasets that can be updated by the research community; data sets include genomes, expressed sequence tags, RNA-Seq, comparative genomics, polymorphism, Chip-Seq, interactome data, gene and protein structure information, images, and etc..

6.4 Closing Remarks

As the human population increases, there is a greater demand for access to the basic necessities of life, including food. At the same time, sustainable agricultural production of food crops is being threatened by diseases, both abiotic and biotic (i.e. pathogens), and has resulted in a decline in the production of crops such as wheat (Lobell *et al.* 2011). In Canada, one of the dominant food crops is wheat, and we are observing an emergence of new plant pathogens that are endangering our production systems (Kelly *et al.* 2015; Ward *et al.* 2008). The Fusarium Head Blight disease of wheat is increasing in incidence, and the population structure of the pathogens is changing, favouring more aggressive pathogens that threaten yields and produce different toxins that contaminate the grains (Clear and Patrick 2010; Graefenham *et al.* 2014; Kelly *et al.* 2015). As plant pathogens adapt to our growing conditions and agricultural practices, there is a greater need for us to understand these pathogens; consequently, it is imperative that we also adapt our management strategies to reduce or eliminate these emerging threats to our

food supply. We exist in a time where food demand is increasing, but food security is being challenged. Understanding the diversity, evolution, and basic biology of pathogens that threaten our food, both past and present, will allow us to better manage and assess risks that pathogens pose to our current and future crop production systems. An important resource for understanding FHB pathogens has been the available genome sequence of *F. graminearum*; in this study, we have provided additional genomic resources to better understand both historical and emerging FHB pathogens in Canada.

REFERENCES:

- Alexander, N. J., S. P. McCormick, C. Waalwijk, T. van der Lee and R. H. Proctor (2011). "The genetic basis for 3-ADON and 15-ADON trichothecene chemotypes in *Fusarium*." *Fungal Genet Biol* 48(5): 485-495.
- Allen, H. K., J. Donato, H. H. Wang, K. A. Cloud-Hansen, J. Davies and J. Handelsman (2010). "Call of the wild: antibiotic resistance genes in natural environments." *Nat Rev Micro* 8(4): 251-259.
- Anders, S. and W. Huber (2010). "Differential expression analysis for sequence count data." *Genome Biol* 11(10): doi: 10.1186/gb-2010-11-10-r106.
- Assefa, S., T. M. Keane, T. D. Otto, C. Newbold and M. Berriman (2009). "ABACAS: algorithm-based automatic contiguation of assembled sequences." *Bioinformatics* 25(15): 1968-1969.
- Bai, G.-H. (1996). "Variation in *Fusarium graminearum* and cultivar resistance to wheat scab." *Plant Dis* 80(9): 975-979.
- Baltrus, D. A., M. T. Nishimura, A. Romanchuk, J. H. Chang, M. S. Mukhtar, K. Cherkis, J. Roach, S. R. Grant, C. D. Jones and J. L. Dangl (2011). "Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates." *PLoS Pathog* 7(7): doi: 10.1371/journal.ppat.1002132.
- Baxter, L., S. Tripathy, N. Ishaque, N. Boot, A. Cabral, E. Kemen, M. Thines, A. Ah-Fong, R. Anderson, W. Badejoko, P. Bittner-Eddy, J. L. Boore, M. C. Chibucos, M. Coates, P. Dehal, K. Delehaunty, S. Dong, P. Downton, B. Dumas, G. Fabro, C. Fronick, S. I. Fuerstenberg, L. Fulton, E. Gaulin, F. Govers, L. Hughes, S. Humphray, R. H. Y. Jiang, H. Judelson, S. Kamoun, K. Kyung, H. Meijer, P. Minx, P. Morris, J. Nelson, V. Phuntumart, D. Qutob, A. Rehmany, A. Rougon-Cardoso, P. Ryden, T. Torto-Alalibo, D. Studholme, Y. Wang, J. Win, J. Wood, S. W. Clifton, J. Rogers, G. Van den Ackerveken, J. D. G. Jones, J. M. McDowell, J. Beynon and B. M. Tyler (2010). "Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome." *Science* 330(6010): 1549-1551.
- Boenisch, M. J. and W. Schäfer (2011). "*Fusarium graminearum* forms mycotoxin producing infection structures on wheat." *BMC Plant Biol* 11(1): doi: 10.1186/1471-2229-11-110.
- Boutigny, A.-L., T. J. Ward, G. J. Van Coller, B. Flett, S. C. Lamprecht, K. O'Donnell and A. Viljoen (2011). "Analysis of the *Fusarium graminearum* species complex from wheat, barley and maize in South Africa provides evidence of species-specific differences in host preference." *Fungal Genet Biol* 48(9): 914-920.
- Brown, N. A., M. Urban, A. M. L. van de Meene and K. E. Hammond-Kosack (2010). "The infection biology of *Fusarium graminearum*: defining the pathways of spikelet to spikelet colonisation in wheat ears." *Fungal Biol* 114(7): 555-571.

Buerstmayr, H., T. Ban and J. A. Anderson (2009). "QTL mapping and marker-assisted selection for Fusarium head blight resistance in wheat: a review." *Plant Breed* 128(1): doi: 10.1111/j.1439-0523.2008.01550.

Carter, J. P., H. N. Rezanoor, D. Holden, A. E. Desjardins, R. D. Plattner and P. Nicholson (2002). "Variation in pathogenicity associated with the genetic diversity of *Fusarium graminearum*." *Eur J Plant Pathol* 108(6): 573-583.

Clear, R. M. and S. K. Patrick (2000). "Fusarium head blight pathogens isolated from fusarium-damaged kernels of wheat in western Canada, 1993 to 1998." *Can J Plant Pathol* 22(1): 51-60.

Clear, R., and S. Patrick. (2010). *Fusarium Head Blight in Canada*. Canadian Grain Commission. Date Accessed: February 2016. Retrieved from: <https://www.grainscanada.gc.ca/str-rst/fusarium/fhbwc-foc-eng.htm>

Conesa, A., S. Götz, J. M. García-Gómez, J. Terol, M. Talón and M. Robles (2005). "Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research." *Bioinformatics* 21(18): 3674-3676.

Cuomo, C. A., U. Güldener, J. R. Xu, F. Trail, B. G. Turgeon, A. Di Pietro, J. D. Walton, L. J. Ma, S. E. Baker, M. Rep, G. Adam, J. Antoniwi, T. Baldwin, S. Calvo, Y. L. Chang, D. Decaprio, L. R. Gale, S. Gnerre, R. S. Goswami, K. Hammond-Kosack, L. J. Harris, K. Hilburn, J. C. Kennell, S. Kroken, J. K. Magnuson, G. Mannhaupt, E. Mauceli, H. W. Mewes, R. Mitterbauer and G. Muehlbauer (2007). "The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization." *Science* 317 (5843): 1400-1402.

Dao, T. T. H., H. J. M. Linthorst and R. Verpoorte (2011). "Chalcone synthase and its functions in plant resistance." *Phytochem Rev* 10(3): 397-412.

Dash, S., J. Van Hemert, L. Hong, R. P. Wise and J. A. Dickerson (2012). "PLEXdb: gene expression resources for plants and plant pathogens." *Nucleic Acids Res* 40: D1194-D1201.

Delcher, A. L., S. Kasif, R. D. Fleischmann, J. Peterson, O. White and S. L. Salzberg (1999). "Alignment of whole genomes." *Nucleic Acids Res* 27(11): 2369-2376.

Denancé, N., B. Szurek and L. D. Noël (2014). "Emerging functions of nodulin-like proteins in non-nodulating plant species." *Plant Cell Physiol* 55(3): 469-474.

Desjardins, A. E. (2006). *Fusarium mycotoxins: chemistry, genetics, and biology*. American Phytopathological Society (APS Press), USA.

Ding, L., H. Xu, H. Yi, L. Yang, Z. Kong, L. Zhang, S. Xue, H. Jia and Z. Ma (2011). "Resistance to hemi-biotrophic *F. graminearum* infection is associated with coordinated and ordered expression of diverse defense signaling pathways." *PLoS ONE* 6(4): doi: 10.1371/journal.pone.0019008.

Doohan, F. M., D. W. Parry, P. Jenkinson and P. Nicholson (1998). "The use of species-specific PCR-based assays to analyse *Fusarium* ear blight of wheat." *Plant Pathol* 47(2): 197-205.

Edwards, S. G., S. R. Pirgozliev, M. C. Hare and P. Jenkinson (2001). "Quantification of trichothecene-producing *Fusarium* species in harvested grain by competitive PCR to determine efficacies of fungicides against *Fusarium* head blight of winter wheat." *Appl Environ Microbiol* 67(4): 1575-1580.

Egmond, H. P. V. and M. A. Jonker (2004). Worldwide regulations for mycotoxins in food and feed in 2003. Food and Agriculture Organization of the United Nations, Italy.

Feng, J., R. Hwang, K. F. Chang, S. F. Hwang, S. E. Strelkov, B. D. Gossen, R. L. Conner and G. D. Turnbull (2010). "Genetic variation in *Fusarium avenaceum* causing root rot on field pea." *Plant Pathol* 59(5): 845-852.

Foroud, N. A., S. P. McCormick, T. MacMillan, A. Badea, D. F. Kendra, B. E. Ellis and F. Eudes (2012). "Greenhouse studies reveal increased aggressiveness of emergent Canadian *Fusarium graminearum* chemotypes in wheat." *Plant Dis* 96(9): 1271-1279.

Frandsen, R., J. Andersson, M. Kristensen and H. Giese (2008). "Efficient four fragment cloning for the construction of vectors for targeted gene replacement in filamentous fungi." *BMC Mol Biol* 9(1): doi: 10.1186/1471-2199-9-70.

Gaffoor, I., D. W. Brown, R. Plattner, R. H. Proctor, W. H. Qi and F. Trail (2005). "Functional analysis of the polyketide synthase genes in the filamentous fungus *Gibberella zeae* (anamorph *Fusarium graminearum*)." *Eukaryot Cell* 4(11): 1926-1933.

Gan, P., K. Ikeda, H. Irieda, M. Narusaka, R. J. O'Connell, Y. Narusaka, Y. Takano, Y. Kubo and K. Shirasu (2013). "Comparative genomic and transcriptomic analyses reveal the hemibiotrophic stage shift of *Colletotrichum* fungi." *New Phytol* 197(4): 1236-1249.

Gardiner, D. M., K. Kazan and J. M. Manners (2009). "Novel genes of *Fusarium graminearum* that negatively regulate deoxynivalenol production and virulence." *Mol Plant Microbe Interact* 22(12): 1588-1600.

Gardiner, D. M., M. C. McDonald, L. Covarelli, P. S. Solomon, A. G. Rusu, M. Marshall, K. Kazan, S. Chakraborty, B. A. McDonald and J. M. Manners (2012). "Comparative pathogenomics reveals horizontally acquired novel virulence genes in fungi infecting cereal hosts." *PLoS Pathog* 8(9): doi: 10.1371/journal.ppat.1002952.

Gardiner, D. M., J. Stiller and K. Kazan (2014). "Genome sequence of *Fusarium graminearum* isolate CS3005." *Genome Announc* 2(2): doi: 10.1128/genomeA.00227-14.

Geiser, D. M., T. Aoki, C. W. Bacon, S. E. Baker, M. K. Bhattacharyya, M. E. Brandt, D. W. Brown, L. W. Burgess, S. Chulze, J. J. Coleman, J. C. Correll, S. F. Covert, P. W. Crous, C. A. Cuomo, G. S. De Hoog, A. Di Pietro, W. H. Elmer, L. Epstein, R. J. N. Frandsen, S. Freeman, T.

Gagkaeva, A. E. Glenn, T. R. Gordon, N. F. Gregory, K. E. Hammond-Kosack, L. E. Hanson, M. d. M. Jimenez-Gasco, S. Kang, H. C. Kistler, G. A. Kuldau, J. F. Leslie, A. Logrieco, G. Lu, E. Lysøe, L.-J. Ma, S. P. McCormick, Q. Migheli, A. Moretti, F. Munaut, K. O'Donnell, L. Pfenning, R. C. Ploetz, R. H. Proctor, S. A. Rehner, V. A. R. G. Robert, A. P. Rooney, B. bin Salleh, M. M. Scandiani, J. Scauflaire, D. P. G. Short, E. Steenkamp, H. Suga, B. A. Summerell, D. A. Sutton, U. Thrane, F. Trail, A. Van Diepeningen, H. D. VanEtten, A. Viljoen, C. Waalwijk, T. J. Ward, M. J. Wingfield, J.-R. Xu, X.-B. Yang, T. Yli-Mattila and N. Zhang (2013). "One fungus, one name: defining the genus *Fusarium* in a scientifically robust way that preserves longstanding use." *Phytopathology* 103(5): 400-408.

Gilbert, J. and A. Tekauz (2000). "Review: recent developments in research on *Fusarium* head blight of wheat in Canada." *Can J Plant Pathol* 22(1): 1-8.

Gilbert, J., D. Abramson, B. McCallum and R. Clear (2002). "Comparison of Canadian *Fusarium graminearum* isolates for aggressiveness, vegetative compatibility, and production of ergosterol and mycotoxins." *Mycopathologia* 153(4): 209-215.

Gilbert, J. and S. Haber (2013). "Overview of some recent research developments in *Fusarium* head blight of wheat." *Can J Plant Pathol* 35(2): 149-174.

Gilbert, J., A. Brûlé-Babel, A. T. Guerrieri, R. M. Clear, S. Patrick, K. Slusarenko and C. Wolfe (2014). "Ratio of 3-ADON and 15-ADON isolates of *Fusarium graminearum* recovered from wheat kernels in Manitoba from 2008 to 2012." *Can J Plant Pathol* 36(1): 54-63.

Glass, N. L. and I. Kaneko (2003). "Fatal attraction: nonself recognition and heterokaryon incompatibility in filamentous fungi." *Eukaryot Cell* 2(1): 1-8.

Goffeau, A., B. Barrell, H. Bussey, R. Davis, B. Dujon, H. Feldmann, F. Galibert, J. Hoheisel, C. Jacq and M. Johnston (1996). "Life with 6000 genes." *Science* 274(5287): 546-567.

Government of Canada. (2010). Market Outlook Report Canada Wheat Sector Profile Part One: Overview. Agriculture and Agri-Food Canada, Canada.

Graefenham, T., S. Patrick, and S. Tittlemier. (2014). *Fusarium* head blight in Canada, maps and charts. Canadian Grain Commission. Date Accessed: February 2016. Retrieved from: <https://www.grainscanada.gc.ca/str-rst/fusarium/fhbmc-feccg-en.htm>

Guo, L., L. Han, L. Yang, H. Zeng, D. Fan, Y. Zhu, Y. Feng, G. Wang, C. Peng, X. Jiang, D. Zhou, P. Ni, C. Liang, L. Liu, J. Wang, C. Mao, X. Fang, M. Peng and J. Huang (2014). "Genome and transcriptome analysis of the fungal pathogen *Fusarium oxysporum* f. sp. *cubense* causing banana vascular wilt disease." *PLoS ONE* 9(4): doi: 10.1371/journal.pone.0095543.

Guo, L., G. Zhao, J.-R. Xu, H. C. Kistler, L. Gao and L.-J. Ma (2016). "Compartmentalized gene regulatory network of the pathogenic fungus *Fusarium graminearum*." *New Phytol* 211(2): 527-541.

- Hansen, F. T., D. M. Gardiner, E. Lysøe, P. R. Fuertes, B. Tudzynski, P. Wiemann, T. E. Sondergaard, H. Giese, D. E. Brodersen and J. L. Sørensen (2015). "An update to polyketide synthase and non-ribosomal synthetase genes and nomenclature in *Fusarium*." *Fungal Genet Biol* 75: 20-29.
- Harris, L. J. and S. C. Gleddie (2001). "A modified Rpl3 gene from rice confers tolerance of the *Fusarium graminearum* mycotoxin deoxynivalenol to transgenic tobacco." *Physiol Mol Plant Pathol* 58(4): 173-181.
- Harris, L. J., M. Balcerzak, A. Johnston, D. Schneiderman and T. Ouellet (2016). "Host-preferential *Fusarium graminearum* gene expression during infection of wheat, barley, and maize." *Fungal Biol* 120(1): 111-123.
- Hindson, C. M., J. R. Chevillet, H. A. Briggs, E. N. Gallichotte, I. K. Ruf, B. J. Hindson, R. L. Vessella and M. Tewari (2013). "Absolute quantification by droplet digital PCR versus analog real-time PCR." *Nat Meth* 10(10): 1003-1005.
- Hou, Z., C. Xue, Y. Peng, T. Katan, H. C. Kistler and J.-R. Xu (2002). "A mitogen-activated protein kinase gene (MGV1) in *Fusarium graminearum* is required for female fertility, heterokaryon formation, and plant infection." *Mol Plant Microbe Interact* 15(11): 1119-1127.
- Jansen, C., D. von Wettstein, W. Schäfer, K.-H. Kogel, A. Felk and F. J. Maier (2005). "Infection patterns in barley and wheat spikes inoculated with wild-type and trichodiene synthase gene disrupted *Fusarium graminearum*." *Proc Natl Acad Sci U S A* 102(46): 16892-16897.
- Jia, L.-J. and W.-H. Tang (2015). "The omics era of *Fusarium graminearum*: opportunities and challenges." *New Phytol* 207(1): 1-3.
- Jones, J. D. G. and J. L. Dangl (2006). "The plant immune system." *Nature* 444(7117): 323-329.
- Kelly, A. C., R. M. Clear, K. O'Donnell, S. McCormick, T. K. Turkington, A. Tekauz, J. Gilbert, H. C. Kistler, M. Busman and T. J. Ward (2015). "Diversity of *Fusarium* head blight populations and trichothecene toxin types reveals regional differences in pathogen composition and temporal dynamics." *Fungal Genet Biol* 82: 22-31.
- Kimura, M., T. Tokai, N. Takahashi-Ando, S. Ohsato and M. Fujimura (2007). "Molecular and genetic studies of *Fusarium* trichothecene biosynthesis: pathways, genes, and evolution." *Biosci Biotechnol Biochem* 71(9): 2105-2123.
- King, R., M. Urban, M. C. U. Hammond-Kosack, K. Hassani-Pak and K. E. Hammond-Kosack (2015). "The completed genome sequence of the pathogenic ascomycete fungus *Fusarium graminearum*." *BMC Genomics* 16(1): doi: 10.1186/s12864-015-1756-1.
- Klosterman, S. J., K. V. Subbarao, S. Kang, P. Veronese, S. E. Gold, B. P. H. J. Thomma, Z. Chen, B. Henrissat, Y.-H. Lee, J. Park, M. D. Garcia-Pedrajas, D. J. Barbara, A. Anchieta, R. de

Jonge, P. Santhanam, K. Maruthachalam, Z. Atallah, S. G. Amyotte, Z. Paz, P. Inderbitzin, R. J. Hayes, D. I. Heiman, S. Young, Q. Zeng, R. Engels, J. Galagan, C. A. Cuomo, K. F. Dobinson and L.-J. Ma (2011). "Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens." PLoS Pathog 7(7): doi: 10.1371/journal.ppat.1002137.

Kristensen, R., M. Torp, B. Kosiak and A. Holst-Jensen (2005). "Phylogeny and toxigenic potential is correlated in *Fusarium* species as revealed by partial translation elongation factor 1 alpha gene sequences." Mycol Res 109(2):173-186.

Kurtz, S., A. Phillippy, A. L. Delcher, M. Smoot, M. Shumway, C. Antonescu and S. L. Salzberg (2004). "Versatile and open software for comparing large genomes." Genome Biol 5(2): doi: 10.1186/gb-2004-5-2-r12.

Laing, C., C. Buchanan, E. N. Taboada, Y. Zhang, A. Kropinski, A. Villegas, J. E. Thomas and V. P. Gannon (2010). "Pan-genome sequence analysis using Panseq: an online tool for the rapid analysis of core and accessory genomic regions." BMC Bioinformatics 11(1): doi: 10.1186/1471-2105-11-461.

Lawler, K., K. Hammond-Kosack, A. Brazma and R. M. Coulson (2013). "Genomic clustering and co-regulation of transcriptional networks in the pathogenic fungus *Fusarium graminearum*." BMC Syst Biol 7(1): doi: 10.1186/1752-0509-7-52.

Lee, T., Y.-K. Han, K.-H. Kim, S.-H. Yun and Y.-W. Lee (2002). "Tri13 and Tri7 determine deoxynivalenol and nivalenol producing chemotypes of *Gibberella zeae*." Appl Environ Microbiol 68(5): 2148-2154.

Lee, J., I.-Y. Chang, H. Kim, S.-H. Yun, J. F. Leslie and Y.-W. Lee (2009). "Genetic diversity and fitness of *Fusarium graminearum* populations from rice in Korea." Appl Environ Microbiol 75(10): 3289-3295.

Leslie, J., and B. Summerell. (2006). The *Fusarium* Laboratory Manual. Blackwell Publishing, USA.

Liang, J., L. Lofgren, Z. Ma, T. J. Ward and H. C. Kistler (2015). "Population subdivision of *Fusarium graminearum* from barley and wheat in the Upper Midwestern United States at the turn of the century." Phytopathology 105(11): 1466-1474.

Lobell, D. B., W. Schlenker and J. Costa-Roberts (2011). "Climate trends and global crop production since 1980." Science 333(6042): 616-620.

Logrieco, A., A. Rizzo, R. Ferracane and A. Ritieni (2002). "Occurrence of beauvericin and enniatins in wheat affected by *Fusarium avenaceum* head blight." Appl Environ Microbiol 68(1): 82-85.

Lysøe, E., K.-Y. Seong and H. C. Kistler (2011). "The transcriptome of *Fusarium graminearum* during the infection of wheat." Mol Plant Microbe Interact 24(9): 995-1000.

Lysøe, E., L. J. Harris, S. Walkowiak, R. Subramaniam, H. H. Divon, E. S. Riiser, C. Llorens, T. Gabaldón, H. C. Kistler, W. Jonkers, A.-K. Kolseth, K. F. Nielsen, U. Thrane and R. J. N. Frandsen (2014). "The genome of the generalist plant pathogen *Fusarium avenaceum* is enriched with genes involved in redox, signaling and secondary metabolism." PLoS ONE 9(11): doi: 10.1371/journal.pone.0112703.

Lysøe, E., R. J. N. Frandsen, H. H. Divon, V. Terzi, L. Orrù, A. Lamontanara, A.-K. Kolseth, K. F. Nielsen and U. Thrane (2016). "Draft genome sequence and chemical profiling of *Fusarium langsethiae*, an emerging producer of type A trichothecenes." Int J Food Microbiol 221: 29-36.

Ma, L.-J., H. C. van der Does, K. A. Borkovich, J. J. Coleman, M.-J. Daboussi, A. Di Pietro, M. Dufresne, M. Freitag, M. Grabherr, B. Henrissat, P. M. Housterman, S. Kang, W.-B. Shim, C. Woloshuk, X. Xie, J.-R. Xu, J. Antoniw, S. E. Baker, B. H. Bluhm, A. Breakspear, D. W. Brown, R. A. E. Butchko, S. Chapman, R. Coulson, P. M. Coutinho, E. G. J. Danchin, A. Diener, L. R. Gale, D. M. Gardiner, S. Goff, K. E. Hammond-Kosack, K. Hilburn, A. Hua-Van, W. Jonkers, K. Kazan, C. D. Kodira, M. Koehrsen, L. Kumar, Y.-H. Lee, L. Li, J. M. Manners, D. Miranda-Saavedra, M. Mukherjee, G. Park, J. Park, S.-Y. Park, R. H. Proctor, A. Regev, M. C. Ruiz-Roldan, D. Sain, S. Sakthikumar, S. Sykes, D. C. Schwartz, B. G. Turgeon, I. Wapinski, O. Yoder, S. Young, Q. Zeng, S. Zhou, J. Galagan, C. A. Cuomo, H. C. Kistler and M. Rep (2010). "Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*." Nature 464(7287): 367-373.

Ma, L.-J., T. Shea, S. Young, Q. Zeng and H. C. Kistler (2014). "Genome sequence of *Fusarium oxysporum* f. sp. *melonis* strain NRRL 26406, a fungus causing wilt disease on melon." Genome Announc 2(4). doi: 10.1128/genomeA.00730-14.

Mayans, O., M. Scott, I. Connerton, T. Gravesen, J. Benen, J. Visser, R. Pickersgill and J. Jenkins (1997). "Two crystal structures of pectin lyase A from *Aspergillus* reveal a pH driven conformational change and striking divergence in the substrate-binding clefts of pectin and pectate lyases." Structure 5(5): 677-689.

Mayer, K. F. X., J. Rogers, J. Doležal, C. Pozniak, K. Eversole, C. Feuillet, B. Gill, B. Friebe, A. J. Lukaszewski, P. Sourdille, T. R. Endo, M. Kubaláková, J. Číhalíková, Z. Dubská, J. Vrána, R. Šperková, H. Šimková, M. Febrer, L. Clissold, K. McLay, K. Singh, P. Chhuneja, N. K. Singh, J. Khurana, E. Akhunov, F. Choulet, A. Alberti, V. Barbe, P. Wincker, H. Kanamori, F. Kobayashi, T. Itoh, T. Matsumoto, H. Sakai, T. Tanaka, J. Wu, Y. Ogihara, H. Handa, P. R. MacLachlan, A. Sharpe, D. Klassen, D. Edwards, J. Batley, O.-A. Olsen, S. R. Sandve, S. Lien, B. Steuernagel, B. Wulff, M. Caccamo, S. Ayling, R. H. Ramirez-Gonzalez, B. J. Clavijo, J. Wright, M. Pfeifer, M. Spannagl, M. M. Martis, M. Mascher, J. Chapman, J. A. Poland, U. Scholz, K. Barry, R. Waugh, D. S. Rokhsar, G. J. Muehlbauer, N. Stein, H. Gundlach, M. Zytynski, V. Jamilloux, H. Quesneville, T. Wicker, P. Faccioli, M. Colaiacovo, A. M. Stanca, H. Budak, L. Cattivelli, N. Glover, L. Pingault, E. Paux, S. Sharma, R. Appels, M. Bellgard, B. Chapman, T. Nussbaumer, K. C. Bader, H. Rimbart, S. Wang, R. Knox, A. Kilian, M. Alaux, F. Alfama, L. Couderc, N. Guillot, C. Viseux, M. Loaec, B. Keller and S. Praud (2014). "A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome." Science 345(6194): doi:

10.1126/science.1251788.

McMullen, M., R. Jones and D. Gallenberg (1997). "Scab of wheat and barley: a re-emerging disease of devastating impact." *Plant Dis* 81(12): 1340-1348.

McMullen, M., G. Bergstrom, E. De Wolf, R. Dill-Macky, D. Herselman, G. Shaner and D. Van Sanford (2012). "A unified effort to fight an enemy of wheat and barley: *Fusarium* head blight." *Plant Dis* 96(12): 1712-1728.

Miedaner, T., C. J. R. Cumagun and S. Chakraborty (2008). "Population genetics of three important head blight pathogens *Fusarium graminearum*, *F. pseudograminearum* and *F. culmorum*." *J Phytopathol* 156(3): 129-139.

Milani, N. A., D. P. Lawrence, A. Elizabeth Arnold and H. D. VanEtten (2012). "Origin of pisatin demethylase (PDA) in the genus *Fusarium*." *Fungal Genet Biol* 49(11): 933-942.

Miller, J. D., A. Taylor and R. Greenhalgh (1983). "Production of deoxynivalenol and related compounds in liquid culture by *Fusarium graminearum*." *Can J Microbiol* 29(9): 1171-1178.

Nalim, F. A., W. H. Elmer, R. J. McGovern and D. M. Geiser (2009). "Multilocus phylogenetic diversity of *Fusarium avenaceum* pathogenic on *Lisianthus*." *Phytopathology* 99(4): 462-468.

Nasmith, C. G., S. Walkowiak, L. Wang, W. W. Y. Leung, Y. Gong, A. Johnston, L. J. Harris, D. S. Guttman and R. Subramaniam (2011). "TRI6 is a global transcription regulator in the phytopathogen *Fusarium graminearum*." *PLoS Pathog* 7(9). doi: 10.1371/journal.ppat.1002266.

Nganje, W. E., D. A. Bangsund, F. L. Leistritz, W. W. Wilson and N. M. Tiapo (2004). "Regional economic impacts of *Fusarium* head blight in wheat and barley." *Appl Econ Perspect Policy* 26(3): 332-347.

O'Connell, R. J., M. R. Thon, S. Hacquard, S. G. Amyotte, J. Kleemann, M. F. Torres, U. Damm, E. A. Buiate, L. Epstein, N. Alkan, J. Altmüller, L. Alvarado-Balderrama, C. A. Bauser, C. Becker, B. W. Birren, Z. Chen, J. Choi, J. A. Crouch, J. P. Duvick, M. A. Farman, P. Gan, D. Heiman, B. Henrissat, R. J. Howard, M. Kabbage, C. Koch, B. Kracher, Y. Kubo, A. D. Law, M.-H. Lebrun, Y.-H. Lee, I. Miyara, N. Moore, U. Neumann, K. Nordstrom, D. G. Panaccione, R. Panstruga, M. Place, R. H. Proctor, D. Prusky, G. Rech, R. Reinhardt, J. A. Rollins, S. Rounsley, C. L. Schardl, D. C. Schwartz, N. Shenoy, K. Shirasu, U. R. Sikhakolli, K. Stuber, S. A. Sukno, J. A. Sweigard, Y. Takano, H. Takahara, F. Trail, H. C. van der Does, L. M. Voll, I. Will, S. Young, Q. Zeng, J. Zhang, S. Zhou, M. B. Dickman, P. Schulze-Lefert, E. Ver Loren van Themaat, L.-J. Ma and L. J. Vaillancourt (2012). "Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses." *Nat Genet* 44(9): 1060-1065.

O'Donnell, K., H. C. Kistler and H. H. Casper (2000). "Gene genealogies reveal global phylogeographic structure and reproductive isolation among lineages of *Fusarium graminearum*, the fungus causing wheat scab." *Proc Natl Acad Sci U S A* 97(14): 7905-7910.

O'Donnell, K., T. J. Ward, D. M. Geiser, H. Corby Kistler and T. Aoki (2004). "Genealogical concordance between the mating type locus and seven other nuclear genes supports formal recognition of nine phylogenetically distinct species within the *Fusarium graminearum* clade." Fungal Genet Biol 41(6): 600-623.

O'Donnell, K., T. J. Ward, D. Abera, H. C. Kistler, T. Aoki, N. Orwig, M. Kimura, Å. Bjørnstad and S. S. Klemsdal (2008). "Multilocus genotyping and molecular phylogenetics resolve a novel head blight pathogen within the *Fusarium graminearum* species complex from Ethiopia." Fungal Genetics and Biology 45(11): 1514-1522.

Otto, T. D., G. P. Dillon, W. S. Degrave and M. Berriman (2011). "RATT: Rapid Annotation Transfer Tool." Nucleic Acids Res 39(9): doi: 10.1093/nar/gkq1268.

Paoletti, M., S. J. Saupe and C. Clavé (2007). "Genesis of a fungal non-self recognition repertoire." PLoS ONE 2(3): doi: 10.1371/journal.pone.0000283.

Picot, A., D. Hourcade-Marcolla, C. Barreau, L. Pinson-Gadais, D. Caron, F. Richard-Forget and C. Lannou (2012). "Interactions between *Fusarium verticillioides* and *Fusarium graminearum* in maize ears and consequences for fungal development and mycotoxin accumulation." Plant Pathol 61(1): 140-151.

Proctor, R. H., S. P. McCormick, N. J. Alexander and A. E. Desjardins (2009). "Evidence that a secondary metabolic biosynthetic gene cluster has grown by gene relocation during evolution of the filamentous fungus *Fusarium*." Mol Microbiol 74(5):1128-1142.

Proctor, R. H., F. Van Hove, A. Susca, G. Stea, M. Busman, T. van der Lee, C. Waalwijk, A. Moretti and T. J. Ward (2013). "Birth, death and horizontal transfer of the fumonisin biosynthetic gene cluster during the evolutionary diversification of *Fusarium*." Mol Microbiol 90(2): 290-306.

Puri, K. D. and S. Zhong (2010). "The 3ADON population of *Fusarium graminearum* found in North Dakota is more aggressive and produces a higher level of DON than the prevalent 15ADON population in spring wheat." Phytopathology 100(10): 1007-1014.

Rudd, J. C., R. D. Horsley, A. L. McKendry and E. M. Elias (2001). "Host plant resistance genes for *Fusarium* head blight." Crop Sci 41(3): 620-627.

Savard, M. E., and B. A. Blackwell. 1994. Spectral Characteristics of Secondary Metabolites from *Fusarium* Fungi. Pages 59-260 in: Mycotoxins in Grain. J.D. Miller, and H.L. Trenholm, eds. Eagan Press, USA.

Schirawski, J., G. Mannhaupt, K. Münch, T. Brefort, K. Schipper, G. Doehlemann, M. Di Stasio, N. Rössel, A. Mendoza-Mendoza, D. Pester, O. Müller, B. Winterberg, E. Meyer, H. Ghareeb, T. Wollenberg, M. Münsterkötter, P. Wong, M. Walter, E. Stukenbrock, U. Guldener and R.

Kahmann (2010). "Pathogenicity determinants in smut fungi revealed by genome comparison." *Science* 330(6010): 1546-1548.

Schmidt, S. M., J. Lukasiewicz, R. Farrer, P. van Dam, C. Bertoldo and M. Rep (2016). "Comparative genomics of *Fusarium oxysporum* f. sp. *melonis* reveals the secreted protein recognized by the Fom-2 resistance gene in melon." *New Phytol* 209(1): 307-318.

Schreiber, K. J., C. G. Nasmith, G. Allard, J. Singh, R. Subramaniam and D. Desveaux (2011). "Found in translation: high-throughput chemical screening in *Arabidopsis thaliana* identifies small molecules that reduce Fusarium head blight disease in wheat." *Mol Plant Microbe Interact* 24(6): 640-648.

Sieber, C. M. K., W. Lee, P. Wong, M. Münsterkötter, H.-W. Mewes, C. Schmeitzl, E. Varga, F. Berthiller, G. Adam and U. Güldener (2014). "The *Fusarium graminearum* genome reveals more secondary metabolite gene clusters and hints of horizontal gene transfer." *PLoS ONE* 9(10): doi: 10.1371/journal.pone.0110311.

Sinha, R. C., M. E. Savard and R. Lau (1995). "Production of monoclonal antibodies for the specific detection of deoxynivalenol and 15-acetyldeoxynivalenol by ELISA." *J Agric Food Chem* 43(6): 1740-1744.

Soanes, D. M., I. Alam, M. Cornell, H. M. Wong, C. Hedeler, N. W. Paton, M. Rattray, S. J. Hubbard, S. G. Oliver and N. J. Talbot (2008). "Comparative genome analysis of filamentous fungi reveals gene family expansions associated with fungal pathogenesis." *PLoS ONE* 3(6): doi: 10.1371/journal.pone.0002300.

Son, H., Y.-S. Seo, K. Min, A. R. Park, J. Lee, J.-M. Jin, Y. Lin, P. Cao, S.-Y. Hong, E.-K. Kim, S.-H. Lee, A. Cho, S. Lee, M.-G. Kim, Y. Kim, J.-E. Kim, J.-C. Kim, G. J. Choi, S.-H. Yun, J. Y. Lim, M. Kim, Y.-H. Lee, Y.-D. Choi and Y.-W. Lee (2011). "A phenome-based functional analysis of transcription factors in the cereal head blight fungus, *Fusarium graminearum*." *PLoS Pathog* 7(10): doi: 10.1371/journal.ppat.1002310.

Sørensen, J. L., T. E. Sondergaard, L. Covarelli, P. R. Fuertes, F. T. Hansen, R. J. N. Frandsen, W. Saei, M. B. Lukassen, R. Wimmer, K. F. Nielsen, D. M. Gardiner and H. Giese (2014). "Identification of the biosynthetic gene clusters for the lipopeptides fusaristatin A and W493 B in *Fusarium graminearum* and *F. pseudograminearum*." *J Nat Prod* 77(12): 2619-2625.

Sperschneider, J., P. N. Dodds, D. M. Gardiner, J. M. Manners, K. B. Singh and J. M. Taylor (2015). "Advances and challenges in computational prediction of effectors from plant pathogenic fungi." *PLoS Pathog* 11(5): doi: 10.1371/journal.ppat.1004806.

Stanke, M. and S. Waack (2003). "Gene prediction with a hidden Markov model and new intron submodel." *Bioinformatics* 19(suppl 2): ii215-ii225.

Starkey, D. E., T. J. Ward, T. Aoki, L. R. Gale, H. C. Kistler, D. M. Geiser, H. Suga, B. Tóth, J.

Varga and K. O'Donnell (2007). "Global molecular surveillance reveals novel *Fusarium* head blight species and trichothecene toxin diversity." *Fungal Genet Biol* 44(11): 1191-1204.

Subramaniam, R., S. Narayanan, S. Walkowiak, L. Wang, M. Joshi, H. Rocheleau, T. Ouellet and L. J. Harris (2015). "Leucine metabolism regulates TRI6 expression and affects deoxynivalenol production and virulence in *Fusarium graminearum*." *Mol Microbiol* 98(4): 760-769.

Swain, M., I. Tsai, S. Assefa, C. Newbold, M. Berriman and T. Otto (2012). "A post-assembly genome-improvement toolkit (PAGIT) to obtain annotated genomes from contigs." *Nat Protoc* 7(7): 1260-1284.

Talas, F. and B. A. McDonald (2015). "Genome-wide analysis of *Fusarium graminearum* field populations reveals hotspots of recombination." *BMC Genomics* 16(1): doi: 10.1186/s12864-015-2166-0.

Talas, F., R. Kalih, T. Miedaner and B. A. McDonald (2016). "Genome-wide association study identifies novel candidate genes for aggressiveness, deoxynivalenol production, and azole sensitivity in natural field populations of *Fusarium graminearum*." *Mol Plant Microbe Interact* 29(5): 417-430.

Trail, F. (2009). "For blighted waves of grain: *Fusarium graminearum* in the postgenomics era." *Plant Physiol* 149(1): 103-110.

Urban, M., R. Pant, A. Raghunath, A. G. Irvine, H. Pedro and K. E. Hammond-Kosack (2014). "The pathogen-host interactions database (PHI-base): additions and future developments." *Nucleic Acids Res* 43:D645-655.

van der Lee, T., H. Zhang, A. van Diepeningen and C. Waalwijk (2015). "Biogeography of *Fusarium graminearum* species complex and chemotypes: a review." *Food Addit Contam* 32(4): 453-460.

Varga, E., G. Wiesenberger, C. Hametner, T. J. Ward, Y. Dong, D. Schöfbeck, S. McCormick, K. Broz, R. Stückler, R. Schuhmacher, R. Krska, H. C. Kistler, F. Berthiller and G. Adam (2015). "New tricks of an old enemy: isolates of *Fusarium graminearum* produce a type A trichothecene mycotoxin." *Environ Microbiol* 17(8): 2588-2600.

Vesonder, R. F., J. J. Ellis, W. F. Kwolek and D. J. DeMarini (1982). "Production of vomitoxin on corn by *Fusarium graminearum* NRRL 5883 and *Fusarium roseum* NRRL 6101." *Appl Environ Microbiol* 43(4): 967-970.

Von der Ohe, C. and T. Miedaner (2011). "Competitive aggressiveness in binary mixtures of *Fusarium graminearum* and *F. culmorum* isolates inoculated on spring wheat with highly effective resistance QTL." *J Phytopathol* 159(6): 401-410.

Wagacha, J. M., E.-C. Oerke, H.-W. Dehne and U. Steiner (2012). "Interactions of *Fusarium* species during prepenetration development." *Fungal Biol* 116(7): 836-847.

Walkowiak, S. and R. Subramaniam (2014). "A nitrogen-responsive gene affects virulence in *Fusarium graminearum*." *Can J Plant Pathol* 36(2): 224-234.

Walkowiak, S., C. T. Bonner, L. Wang, B. Blackwell, O. Rowland and R. Subramaniam (2015). "Intraspecies interaction of *Fusarium graminearum* contributes to reduced toxin production and virulence." *Mol Plant Microbe Interact* 28(11): 1256-1267.

Walter, S., P. Nicholson and F. M. Doohan (2010). "Action and reaction of host and pathogen during *Fusarium* head blight disease." *New Phytol* 185(1): 54-66.

Wang, J.-H., M. Ndoye, J.-B. Zhang, H.-P. Li and Y.-C. Liao (2011a). "Population structure and genetic diversity of the *Fusarium graminearum* species complex." *Toxins* 3(8): 1020-1037.

Wang, C., S. Zhang, R. Hou, Z. Zhao, Q. Zheng, Q. Xu, D. Zheng, G. Wang, H. Liu, X. Gao, J.-W. Ma, H. C. Kistler, Z. Kang and J.-R. Xu (2011b). "Functional analysis of the kinome of the wheat scab fungus *Fusarium graminearum*." *PLoS Pathog* 7(12): doi: 10.1371/journal.ppat.1002460.

Wang, L., C. Mogg, S. Walkowiak, M. Joshi and R. Subramaniam (2014). "Characterization of NADPH oxidase genes NoxA and NoxB in *Fusarium graminearum*." *Can J Plant Pathol* 36(1): 12-21.

Ward, T. J., J. P. Bielawski, H. C. Kistler, E. Sullivan and K. O'Donnell (2002). "Ancestral polymorphism and adaptive evolution in the trichothecene mycotoxin gene cluster of phytopathogenic *Fusarium*." *Proc Natl Acad Sci U S A* 99(14): 9278-9283.

Ward, T. J., R. M. Clear, A. P. Rooney, K. O'Donnell, D. Gaba, S. Patrick, D. E. Starkey, J. Gilbert, D. M. Geiser and T. W. Nowicki (2008). "An adaptive evolutionary shift in *Fusarium* head blight pathogen populations is driving the rapid spread of more *toxigenic Fusarium graminearum* in North America." *Fungal Genet Biol* 45(4): 473-484.

Wiemann, P., C. M. K. Sieber, K. W. von Bargaen, L. Studt, E.-M. Niehaus, J. J. Espino, K. Huß, C. B. Michielse, S. Albermann, D. Wagner, S. V. Bergner, L. R. Connolly, A. Fischer, G. Reuter, K. Kleigrewe, T. Bald, B. D. Wingfield, R. Ophir, S. Freeman, M. Hippler, K. M. Smith, D. W. Brown, R. H. Proctor, M. Münsterkötter, M. Freitag, H.-U. Humpf, U. Güldener and B. Tudzynski (2013). "Deciphering the cryptic genome: genome-wide analyses of the rice pathogen *Fusarium fujikuroi* reveal complex regulation of secondary metabolism and novel metabolites." *PLoS Pathog* 9(6): doi: 10.1371/journal.ppat.1003475.

Windels, C. E. (2000). "Economic and social impacts of *Fusarium* head blight: changing farms and rural communities in the northern great plains." *Phytopathology* 90(1): 17-21.

Wong, P., M. Walter, W. Lee, G. Mannhaupt, M. Münsterkötter, H.-W. Mewes, G. Adam and U. Güldener (2011). "FGDB: revisiting the genome annotation of the plant pathogen *Fusarium graminearum*." *Nucleic Acids Res* 39: D637-D639.

World Health Organization, and Food and Agriculture Organization of the United Nations. 2011. Evaluation of Certain Contaminants in Food. World Health Organization (WHO). Switzerland.

Xu, X. M., D. W. Parry, P. Nicholson, M. A. Thomsett, D. Simpson, S. G. Edwards, B. M. Cooke, F. M. Doohan, J. M. Brennan, A. Moretti, G. Tocco, G. Mule, L. Hornok, G. Giczey and J. Tatnell (2005). "Predominance and association of pathogenic fungi causing *Fusarium* ear blight in wheat in four European countries." *Eur J Plant Pathol* 112(2): 143-154.

Xu, J.-R., Y.-L. Peng, M. B. Dickman and A. Sharon (2006). "The dawn of fungal pathogen genomics." *Annu Rev Phytopathol* 44(1): 337-366.

Yang, L., T. van der Lee, X. Yang, D. Yu and C. Waalwijk (2008). "*Fusarium* populations on Chinese barley show a dramatic gradient in mycotoxin profiles." *Phytopathology* 98(6): 719-727.

Yun, Y., Z. Liu, Y. Yin, J. Jiang, Y. Chen, J.-R. Xu and Z. Ma (2015). "Functional analysis of the *Fusarium graminearum* phosphatome." *New Phytol* 207(1): 119-134.

Zhang, H., Z. Zhang, T. van der Lee, W. Q. Chen, J. Xu, J. S. Xu, L. Yang, D. Yu, C. Waalwijk and J. Feng (2010). "Population genetic analyses of *Fusarium asiaticum* populations from barley suggest a recent shift favoring 3ADON producers in Southern China." *Phytopathology* 100(4): 328-336.

Zhao, X.-M., X.-W. Zhang, W.-H. Tang and L. Chen (2009). "FPPI: *Fusarium graminearum* protein-protein interaction database." *J Proteome Res* 8(10): 4714-4721.

Zhao, C., C. Waalwijk, P. J. G. M. de Wit, D. Tang and T. van der Lee (2013). "RNA-Seq analysis reveals new gene models and alternative splicing in the fungal pathogen *Fusarium graminearum*." *BMC Genomics* 14(1): doi: 10.1186/1471-2164-14-21.