Genetic and Pathogenic Differences Between *Microdochium nivale* and *Microdochium majus*

by

Linda Elizabeth Jewell

A Thesis
Presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Doctor of Philosophy
in
Environmental Science

Guelph, Ontario, Canada

© Linda Elizabeth Jewell, December, 2013
GENETIC AND PATHOGENIC DIFFERENCES BETWEEN

MICROODOCHIUM NIVALE AND MICROODOCHIUM MAJUS

Linda Elizabeth Jewell
University of Guelph, 2013

Microdochium nivale and M. majus are fungal plant pathogens that cause cool-temperature diseases on grasses and cereals. Nucleotide sequences of four genetic regions were compared between isolates of M. nivale and M. majus from Triticum aestivum (wheat) collected in North America and Europe and for isolates of M. nivale from turfgrasses from both continents. Draft genome sequences were assembled for two isolates of M. majus and two of M. nivale from wheat and one from turfgrass. Dendograms constructed from these data resolved isolates of M. majus into separate clades by geographic origin. Among M. nivale, isolates were instead resolved by host plant species. Amplification of repetitive regions of DNA from M. nivale isolates collected from two proximate locations across three years grouped isolates by year, rather than by location. The mating-type (MAT1) and associated flanking genes of Microdochium were identified using the genome sequencing data to investigate the potential for these pathogens to produce ascospores. In all of the Microdochium genomes, and in all isolates assessed by PCR, only the MAT1-2-1 gene was identified. However, unpaired, single-conidium-derived colonies of M. majus produced fertile perithecia in the lab. This finding contrasts with the cannonical requirements for sexual spore production among the Ascomycota. To further explore this, MAT1 and flanking gene sequences were identified in the genome sequences of six additional species.
from Xylariaceae, no homologs of known MAT1-1-1 genes were detected, suggesting that the control of sexual reproduction among the Xylariaceae may be differentially regulated relative to other Sordariomycete species. Detached leaves of *T. aestivum* and *Poa pratensis* (Kentucky bluegrass) were inoculated with either *M. nivale* or *M. majus* and were incubated at either 23 °C or at 4 °C to investigate the infection processes of these pathogens. Despite reported field host preferences, the two pathogens were equally virulent on both host plants at the temperatures investigated. The results presented here reveal genetic, but not pathogenic, differences between *M. nivale* and *M. majus* and further demonstrate that sub-populations may exist within the groups of these pathogens on different host plants.
ACKNOWLEDGEMENTS

Firstly, and most importantly, I would like to thank my advisor Dr. Tom Hsiang for accepting me into his lab and inviting me to participate in this exciting and multifaceted research project. I have sincerely appreciated the opportunity that he has given me to learn a very diverse set of skills. I am also very appreciative of the opportunities that he has given me to participate in engaging and inspiring conferences throughout my time as a student. I would also like to thank all of the members of my thesis, qualifying examination, and defense committees for their patience, guidance, and helpful suggestions at every stage of my project. Thanks are also extended to the administrative and support staff at the University of Guelph, especially in the School of Environmental Sciences, for helping to make my time here at the U of G run smoothly. I would also like to thank the numerous collaborators who provided me with some of the samples or materials that contributed to these analyses.

I thank my friends and my family, especially my parents Bernice and Calvin and my brother Chris for their love and support through all of the long years that I have been in school. Thank you to all of my labmates, both past and current, for their friendship and help. I would especially like to thank Amy Shi for her help with research, for her amazing skills as a tour guide, and for her cat-wrangling abilities. Thank you Anne-Miet, for being a morning person; to Mihaela, for support and friendship; to Sarah, for being a vegetarian and a travel buddy; to Vince, for his help with bioinformatics; to Brady, for his help with RNA; to unbelievable undergrads Holly, Sara, and Craig, for their assistance in the lab and their and friendship; and to everyone else I have had the pleasure of working with, for their kindness and assistance.

Finally, because everyone would be horrified if I left them out, thank you Jim and Luke for your skills as alarm clocks and as supervisors of all fridge-related endeavours.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ................................................................................................. iv

TABLE OF CONTENTS .................................................................................................... v

LIST OF TABLES ............................................................................................................... xi

LIST OF FIGURES ........................................................................................................... xvi

LIST OF APPENDICES ................................................................................................... xxiii

LIST OF ABBREVIATIONS AND ACRONYMS ........................................................... xxvi

Chapter 1 General Introduction & Literature Review ................................................... 1

1.1 Introduction ............................................................................................................... 1

1.2 General information about *Microdochium nivale* and *M. majus* .................. 2

1.2.1 Disease Cycle ...................................................................................................... 4

1.2.2 Phylogenetic classification of *M. nivale* and *M. majus* ......................... 5

1.3 Differences between *M. majus* and *M. nivale* ................................................. 6

1.3.1 Morphological characteristics ......................................................................... 7

1.3.2 Pathogenic differences ..................................................................................... 7

1.3.3 Genetic differences ............................................................................................ 8

1.4 Sexual Reproduction .............................................................................................. 9

1.4.1 Sexual reproduction in the Ascomycota ....................................................... 9

1.4.2 Sexual reproduction in *Microdochium nivale* and *M. majus* ............... 11

1.5 Sequencing techniques and bioinformatics ........................................................... 12

1.5.1 DNA sequencing techniques ....................................................................... 12

1.5.2 RNA sequencing ............................................................................................. 14

1.6 Hypotheses and Objectives ............................................................................... 16
Chapter 3 Comparative Genomics ................................................................. 107

3.1 Introduction ............................................................................................. 107

3.1.1 General overview of whole-genome analyses ...................................... 107

3.1.2 Sequencing platforms .......................................................................... 108

3.1.3 Genome assembly and protein prediction ............................................. 111

3.1.4 Whole-genome comparisons ................................................................. 114

3.1.5 Objectives ............................................................................................ 118

3.2 Materials and Methods ........................................................................... 118

3.2.1 DNA extraction, quantification, and sequencing .................................... 118

3.2.2 Genome assembly and gene prediction ............................................... 119

3.2.3 Whole-genome comparisons and identification of unique genes .......... 121

3.2.4 Design of species-specific primers ....................................................... 122

3.2.5 Identification of putative pathogen-related genes ................................ 123

3.2.6 Identification of putative transposable elements .................................. 124

3.3 Results ..................................................................................................... 124

3.3.1 Genome sequencing, assembly, and protein prediction ....................... 124

3.3.2 Whole-genome comparisons ................................................................. 125

3.3.3 Development of species-specific primers ............................................. 127

3.3.4 Identification of pathogenesis-related genes ........................................ 128

3.3.5 Identification of putative transposable elements and comparison to PHI genes . 129

3.3.6 Identification of a putative EF-1α sequence in the genome of M. bolleyi ........ 130

3.4 Discussion ............................................................................................... 131

3.5 References for Chapter 3 ......................................................................... 139

vii
5.1.1 Reproduction in Fungi .......................................................................................... 275
5.1.2 Genes associated with sexual reproduction .................................................... 278
5.1.3 Sexual reproduction in *M. nivale* and *M. majus* ........................................... 280
5.1.4 Objectives ........................................................................................................ 282
5.2 Materials and Methods ......................................................................................... 282
5.2.1 Test of mating type primers based on conserved sequences ......................... 282
5.2.2 Identification of putative MAT1 loci and flanking genes and screening of isolate
collection .................................................................................................................. 283
5.2.3 Mating experiments .......................................................................................... 286
5.2.4 Comparison of *Microdochium* sp. with other species .................................. 287
5.3 Results .................................................................................................................. 288
5.3.1 Test of published mating type primers and redesigned universal primers ....... 288
5.3.2 Identification of putative mating type and flanking genes .............................. 289
5.3.3 Mating experiments ......................................................................................... 293
5.3.4 Comparison of the *Microdochium* MAT1 locus to that of other species ....... 297
5.4 Discussion ............................................................................................................ 298
5.5 References for Chapter 5 ...................................................................................... 312
Appendices for Chapter 5 ......................................................................................... 344
Chapter 6 Infection Process of *Microdochium majus* and *M. nivale* .................... 388
6.1 Introduction ........................................................................................................... 388
6.1.1 The disease cycle ............................................................................................ 388
6.1.2 The infection processes of *M. nivale* and *M. majus* ................................. 389
6.1.3 Host specificity and infection success ............................................................. 391
6.1.4 Objectives .............................................................................................................. 391

6.2 Materials and Methods .......................................................................................... 392

6.2.1 Plant culture ......................................................................................................... 392

6.2.2 Inoculum preparation .......................................................................................... 392

6.2.3 Inoculation ............................................................................................................ 393

6.2.4 Sample collection, staining, and scoring ............................................................ 394

6.2.5 Statistical methods .............................................................................................. 395

6.3 Results ....................................................................................................................... 396

6.3.1 Experiment A ....................................................................................................... 396

6.3.2 Experiment B ....................................................................................................... 397

6.3.3 Experiment C ....................................................................................................... 398

6.3.4 Experiment D ....................................................................................................... 399

6.3.5 Experiment E ....................................................................................................... 400

6.4 Discussion ................................................................................................................ 401

6.5 References for Chapter 6 ....................................................................................... 407

Appendices for Chapter 6 ............................................................................................ 426

Chapter 7 General Discussion and Conclusions ........................................................ 427

7.1 Major conclusions ................................................................................................... 427

7.2 General Discussion and Conclusions ..................................................................... 428

7.3 References for Chapter 7 ....................................................................................... 438
LIST OF TABLES

Table 2.1 Isolates of *Microdochium nivale* and *M. majus*, including their geographic origin and host-plant origin used for nucleotide sequence analysis................................................................. 53

Table 2.2 Primers used in PCR and sequencing reactions............................................................................................................................... 56

Table 2.3 List of RPB2 sequences used to design primers for *M. nivale* and *M. majus* (Sordariomycetes, Xylariales) with taxonomic information and GenBank accession numbers... 57

Table 2.4 List of species used to design β-tubulin primers for *M. nivale* and *M. majus* with GenBank accession numbers. Other than the *Microdochium* species, all species included are members of the Xylariaceae....................................................................................................................... 59

Table 3.1 Summary of DNA quantity and sequencing facility utilized for genome sequencing of six *Microdochium* isolates. ........................................................................................................................................... 145

Table 3.2 Sordariomycetes genomes included in whole-genome comparisons against *Microdochium* spp. ............................................................................................................................................. 146

Table 3.3 Summary of genome assembly, protein prediction, and predicted gene annotation statistics for sequenced *Microdochium* genomes........................................................................................................... 147

Table 3.4 Genome assembly statistics for *M. majus* assembly with Velvet, SOAPdenovo, and ABysS for odd-numbered kmers 29-59. Note that gap-closing was not performed for this comparison........................................................................................................................................... 148

Table 3.5 Species-specific primers designed and tested with at least two isolates each of *M. nivale* and *M. majus* .................................................................................................................................................. 149

Table 3.6 Comparisons between the predicted gene sequences from *M. nivale*, *M. majus*, and *M. bolleyi* against each other and against the genomes of 6 (non-*Microdochium*) members of the
Xylariales and 9 non-Xylariales members of the Sordariomycetes. Comparisons were performed using tBLASTn with an e-value cutoff of 1e-05.

Table 3.7 Comparisons between the predicted gene sequences from *M. nivale*, *M. majus*, and *M. bolleyi* against each other and against the genomes of 6 (non-*Microdochium*) members of the Xylariales and 9 non-Xylariales members of the Sordariomycetes. Comparisons were performed using tBLASTn with an e-value cutoff of 1e-20.

Table 3.8 Comparisons between the predicted gene sequences from *M. nivale*, *M. majus*, and *M. bolleyi* against each other and against the genomes of 6 (non-*Microdochium*) members of the Xylariales and 9 non-Xylariales members of the Sordariomycetes. Comparisons were performed using tBLASTn with an e-value cutoff of 1e-50.

Table 3.9 Fungal pathogen-host interaction (PHI) genes with highly variable copy numbers among *Microdochium* predicted gene sets that may play a role in pathogenicity or fungicide resistance (*Mb = Microdochium bolleyi; Mm = M. majus; Mn = M. nivale*).

Table 3.10 Accession numbers for ten randomly-selected putative unique predicted gene sequences identified in *Microdochium* spp.

Table 3.11 Transposable element sequences downloaded from GenBank that were used in comparisons against the genomes of *Microdochium* spp.

Table 3.12 Summary of putative transposable element sequences identified in the *Microdochium* spp. genomes and their relative proximity to putative pathogen-host interaction genes.

Table 4.1 Year and location of collection from the Guelph Turfgrass Institute for all samples included in multi-year screening. See Figure 4.3 for a map depicting these locations.

Table 4.2 List of all SSR and ISSR primers screened to assess genetic variation in *Microdochium nivale* field isolates collected across three years.
Table 4.3 List of all SSR and ISSR primers selected for analysis of genetic variation in *Microdochium nivale* field isolates collected across three years. .............................................. 262
Table 4.4 Results of linkage disequilibrium calculations performed using the program Disequil (described in Mahuku et al. 1998) on isolates of *Microdochium nivale* collected in three separate years and in two locations at the Guelph Turfgrass Institute. See Figure 4.3 for a map depicting these locations. .............................................................................................................. 263
Table 5.1 Published mating-type primers tested with *M. nivale* and *M. majus* ................. 316
Table 5.2 Isolates of *M. nivale* and *M. majus* tested with published mating-type primers.... 317
Table 5.3 List of species and GenBank accession numbers for genes used to design conserved MAT1-1-1 and MAT1-2-1 primers in *Microdochium* sp. ................................................................. 318
Table 5.4 Primers designed to amplify mating-type (MAT1-1 and MAT1-2) and flanking genes in *Microdochium* spp. ........................................................................................................... 320
Table 5.5 Number of bands amplified by the ISSR and SSR PCR primers listed with a selection of *Microdochium majus* isolates including the single-gamete-derived (AS) cultures derived from the parent isolate 99049. ............................................................................................................. 321
Table 5.6 List of species and GenBank accession numbers for genes used to search for putative MAT1-1-1, MAT1-2-1, and flanking genes in *Microdochium* genomes by standalone tBLASTn ................................................................. 322
Table 5.7 List of predicted genes corresponding to the mating-type (MAT1-2-1) and flanking genes (cytoskeletal assembly protein SLA2 and DNA lyase APN2) in the *Microdochium* genomes studied. Comparisons were performed using standalone tBLASTn to query the gene of interest against the *Microdochium* spp. genomes listed................................................................. 323
Table 5.8 Results of MAT1-2-1 amplification of *M. nivale* isolates with the primers Mn_MAT2_20F and Mn_MAT2_727R ................................................................. 324

Table 5.9 Isolates used in mating-type crosses .......................................................... 325

Table 5.10 Summary of perithecial production in second experiment according to temperature of incubation and species included in each cross. Isolates were inoculated on sterilized wheat straw overlaid on water agar, and observations were performed after two months of incubation. . 326

Table 5.11 Species used for MAT-region synteny investigation. All species were members of the order Xylariales and family Xylariaceae ....................................................... 327

Table 5.12 BLASTx results for putative matches to flanking and mating genes observed in Xylariales genomes. BLAST searches were performed by querying the putative flanking genes (cytoskeletal assembly protein SLA2, DNA lyase APN2, anaphase-promoting complex protein APC5, and cytochrome oxidase COX) and mating type gene (MAT2) against the GenBank non-redundant database ................................................................................................................................. 328

Table 6.1 Summary of conditions tested in infection process experiments performed for *M. nivale* and *M. majus* inoculated on *P. pratensis* and *T. aestivum*. .............................................. 409

Table 6.2 Isolates of *M. nivale* and *M. majus* used in infection process experiments on *T. aestivum* and *P. pratensis*. ........................................................................................................ 410

Table 6.3 Sample collection timepoints for infection process studies of *M. nivale* and *M. majus* on *T. aestivum* and *P. pratensis*. ................................................................................ 411

Table 6.4 Comparisons between mean number of penetration observations per unit area for each isolate of *Microdochium nivale* (Mn) and *M. majus* (Mm) on detached leaves of Kentucky bluegrass (K) and wheat (W) in experiment D. Data for each isolate at each time point were
Table 6.5 Comparisons between mean number of penetration observations per unit area for all isolates, regardless of identity, from each host type on Kentucky bluegrass and wheat in experiment D. Data for each host type at each time point were compared using a two-sided Wilcoxon rank sum approximation. Means followed by different letters were significantly different at \( p = 0.05 \).

Table 6.6 Comparisons between mean number of penetration observations per unit area for each isolate of *Microdochium nivale* (Mn) or *M. majus* (Mm) on detached leaves of Kentucky bluegrass (K) and wheat (W) in experiment E. Data for each isolate at each time point were compared using a two-sided Wilcoxon rank sum approximation. Means followed by different letters were significantly different at \( p = 0.05 \).

Table 6.7 Comparisons between mean number of incidences of penetration per unit area for all isolates from each host type on detached leaves of Kentucky bluegrass and wheat in experiment E. Data for each host type at each time point were compared using a two-sided Wilcoxon rank sum approximation. Means followed by different letters were significantly different at \( p = 0.05 \).
LIST OF FIGURES

Figure 1.1 Pink snow mould damage on mixed creeping bentgrass / annual bluegrass green (Agrostis stolonifera / Poa annua), Guelph, ON, 02 March 2012................................................................. 24

Figure 2.1 Number of research journal papers published between 2000-2012 including the terms "maximum likelihood", "maximum parsimony", and either "neighbor joining" or "neighbour joining" obtained by searching the google scholar database using the term phylog* and either "maximum likelihood", "maximum parsimony" or "'neighbour joining' OR 'neighbor joining'" for articles published between 2000 and 2012. .................................................................................. 60

Figure 2.2 Gel image for European (lanes A-F) and North American (lanes G-I, K-O) isolates of Microdochium nivale amplified with the Y13NF and Y13NR primers, which target a genetic region of unknown function. All six of the European isolates (lanes A-F) were strongly amplified, whereas weak bands are present in only two (lanes H, I) of the North American samples. A negative control, containing water instead of DNA, was included in the reaction (lane Q). Select sizes (bp) from a 100 bp increment ladder (lane J) are indicated to the right............. 61

Figure 2.3 Bootstrapped maximum likelihood trees for RPB2, β-tubulin, EF-1α, and ITS. The tips of the tree are labelled with isolate number, species (M or N for M. majus or M. nivale, respectively), host origin (W or T for wheat or turfgrass), and geographic origin (NA or EU for North America or Europe). Bootstrap values out of 100 are located on the respective branches. Nodes with less than 50% support were collapsed. Separate clades resolving the sister species and host-specific or geographic groupings are indicated by curly braces. ........................................ 63

Figure 2.4 Gel image for HindIII digest of RPB2 amplicon of M. nivale (lanes A-C) and M. majus (lanes D-F). Partial RPB2 sequences were amplified with the primers RPB_150F and fRPB2-7cR. The amplicons of M. nivale isolates were digested at two locations to produce...
fragments that were 125, 444, and 475 bp in length and the *M. majus* amplicons were digested at only one location, producing fragments that were 444 and 596 bp in length. Note that the 444 and 475 bp bands were not well-resolved. A non-digested RPB2 amplicon (1,040 bp in length) from *M. majus* is included for comparison (lane G). Select sizes (bp) from a 100 bp increment ladder (lane J) are indicated to the right.

Figure 3.1 Pipeline for DNA sequencing by Illumina-Solexa technology. Genomic DNA (1) is isolated and (2) sheared into millions of fragments. Adapter sequences are ligated onto both sides of all fragments (3). The fragments are introduced to the solid surface (4), which contains sequences that are complementary to the adapters. In the PCR stage (5), the second adapter sequence of each fragment (white) anneals to its complement. The fragment is amplified by DNA polymerase to produce a complementary strand (6). In each round of PCR, both strands are duplicated (7). The result is clusters of identical complementary sequences (8) for each of the original fragments. In the sequencing stage, a solution containing all four dNTPs, each labelled with a different terminating fluorophore, is introduced (9). Only one nucleotide at a time can be introduced (10) due to the terminator sequence. The incorporated nucleotide is identified by the fluorophore. The fluorophore terminator is cleaved (11) to allow for the addition of the next nucleotide. The labelled dNTP solution is re-applied (12), the second nucleotide is incorporated (13), and step 11 is repeated. This process is repeated until the full fragment has been sequenced (14) (Mardis 2008).

Figure 3.2 Summary of paired-end sequencing. Genomic DNA is sheared into fragments of approximately the same length (e.g. approximately 500 bp). This fragment is then sequenced from both the 5' and the 3' ends, generating two "paired" fragments (labelled A and B and joined by a dashed arc). The number of sequencing cycles performed is equal to the lengths of these..
fragments (e.g. for strands consisting of two paired-end fragments of 100 bp each, a total of 200 cycles would be performed). During the assembly process, the expected distance between these two fragments (in this example, 300 bp) is used to facilitate their association with the other fragments generated during sequencing.

Figure 3.3 Schematic representation of a de Bruijn graph (Zerbino and Birney 2008). Nodes are represented by boxes. Each node consists of a short alignment of overlapping sequences of the same length. Each node also has a sister node consisting of the reverse complement of the sequences and the alignment in its sister, located immediately above or below the node (e.g. A and A' are sister nodes). Nodes are connected based on their sequence similarities: for example, the final sequence in node A shares four of its six nucleotides with the first sequence in node B. Node B is connected both to node A and to nodes C and D, because the final sequence in node B overlaps equally well with the first sequences in both nodes C and D. The use of de Bruijn graphs results in the association of short, overlapping alignments that are used to assemble the sequencing reads into larger contigs or scaffolds.

Figure 3.4 a) Neighbour-joining, b) maximum likelihood, and c) maximum parsimony trees depicting the relationships between the sequenced Microdochium genomes, based on the concatenated sequences of ten genes that were putatively unique to Microdochium. Bootstrap values (out of 100) are displayed on each node. The ten genes are listed in Table 3.10. Scale bars represent either 0.1 nucleotide change per base (a and b) or 100 substitutions (c).

Figure 4.1 Diagram depicting the relative positions of hypothetical SSR and ISSR loci and the primers that could amplify these regions. The SSR loci (white text on black background) are flanked by ISSR regions (black text on white background). Whereas the ISSR primer (white text on black arrow) is anchored within the repetitive SSR region, and thus can be designed with only
knowledge about the repetitive SSR sequence, the SSR primers (black text on white arrows) are located within the non-repetitive ISSR sequences, and thus require more detailed genomic information.

Figure 4.2 Grasses displaying symptoms of pink snow mold and / or Fusarium patch. Both photos were taken at the GTI, Guelph, Ontario, 2 March 2012. a) Kentucky bluegrass in the area to the east of the native green and b) annual bluegrass / creeping bentgrass mixture, native green.

Figure 4.3 Map of the Guelph Turfgrass Institute with collection locations and key landmarks indicated: A: Near pathology green (PG); B: roadside (Rd). The grass species at both collection locations was Poa pratensis (Kentucky bluegrass).

Figure 4.4 UPGMA tree depicting relationships between M. nivale isolates collected from P. pratensis at two locations (Figure 4.3) yearly from 2011-2013. Bootstrap values (out of 100) are displayed on key nodes. Legend: 1: 2013 pathology green isolates; 2: 2013 roadside; 3: 2012 pathology green; 4: 2012 roadside; 5: 2011 pathology green; 6: 2011 roadside.

Figure 5.1 Orientation of the MAT1 region and flanking genes in the Sordariomycete species Neurospora crassa (Butler et al. 2004), Giberella zeae (Yun et al. 2000), Botrytis cinerea, and Sclerotinia sclerotorium (Amselem et al. 2001). Diagram is not to scale. A vertical bar extending over the MAT1 locus indicates that this species is heterothallic and that the gene(s) located on the parallel bars are interchangeable in the two mating types.

Figure 5.2 Orientation and synteny of the putative MAT1 region and the flanking genes APC5, SLA2, APN2, and COX13 in several species of Xylariales, including Microdochium sp. Diagrams are not to scale. A double slash (/\) indicates a long (>10 kb) distance between putative
genes found on the same scaffold, and a vertical bar (|) indicates that the gene(s) that follow was / were found on a different scaffold. .......................... 339

Figure 5.3 Single segment of wheat straw inoculated with *M. majus* isolate 99049 and *M. nivale* isolate 99077 incubated on water agar at 20 °C for approximately two months. Note the production of perithecia on the side closest to *M. majus* (A) but not *M. nivale* (B). 10x magnification. .......................... 340

Figure 5.4 Perithecium of *M. majus* isolate 99049 at (A) 40 x and (B) 100x magnification. Perithecia depicted were observed after two months of incubation on wheat straw on water agar at 20 °C.......................... 341

Figure 5.5 Ascospore (centre) produced by *M. majus* isolate 99049 at 400x magnification. Spore depicted were observed after two months of incubation on wheat straw on water agar at 20 °C. .......................... 342

Figure 5.6 Bootstrapped UPGMA tree depicting the relationships between ten single-ascospore cultures derived from *M. majus* isolate 99049 relative to their parent culture and DNA from seven other *M. majus* isolates, including one isolate collected from the same location on the same date as the parent culture (99061), two isolates from Europe (10098 and 10099), and four cultures collected from the same wheat field on the same date (12043-12046). The horizontal bar represents 10% sequence divergence. Bootstrap values are out of 100.......................... 343

Figure 6.1 The disease cycle describing the events that occur during a host-pathogen interaction (modified from (Agrios 2005)).......................... 416

Figure 6.2 Number of detached leaf segments of *P. pratensis* (A) and *T. aestivum* (B) where penetration was observed at time of collection for leaf blades treated with hyphal inoculum in experiment A. Three leaf blades were collected at each time point. .......................... 417
Figure 6.3 Penetration of stomata of detached leaves of *T. aestevum* by hyphae of *M. majus* isolate 99049 (stained blue) incubated on moistened filter paper and incubated at 22 ºC. Photo taken at 3 dpi, 400 x magnification................................................................. 418

Figure 6.4 Hyphae of *M. majus* isolate 99061 emerging from stomata of detached leaves of wheat (circled) incubated on moistened filter paper and incubated at 22 ºC. Photo taken at 400x magnification, 5 dpi. .................................................................................................................. 419

Figure 6.5 Number of detached leaf segments of *P. pratensis* (A) and *T. aestivum* (B) incubated on moist filter paper where penetration was observed at time of collection for leaf blades treated with hyphal inoculum in experiment B. Three leaf blades were collected at each time point. ... 420

Figure 6.6 Number of detached leaf segments of *P. pratensis* (A) and *T. aestivum* (B) incubated on moist filter paper where penetration was observed at time of collection for leaf blades treated with conidial inoculum in experiment B. Three leaf blades were collected at each time point. 421

Figure 6.7 Number of detached leaf segments of *P. pratensis* (A) and *T. aestivum* (B) incubated on moist filter paper where penetration was observed at time of collection for leaf blades treated with hyphal inoculum in experiment C................................................................. 422

Figure 6.8 Number of detached leaf segments of *P. pratensis* (A) and *T. aestivum* (B) incubated on moist filter paper where penetration was observed at time of collection for leaf blades treated with conidial inoculum in experiment C........................................................................................................ 423

Figure 6.9 Number of incidences of penetration per unit area on *P. pratensis* (A) and *T. aestivum* (B) incubated on moist filter paper at time of collection for leaf blades treated with hyphal inoculum in experiment D. ................................................................................................................. 424
Figure 6.10 Number of incidences of penetration per unit area on *P. pratensis* (A) and *T. aestivum* (B) incubated on moist filter paper at time of collection for leaf blades treated with hyphal inoculum in experiment E. ........................................................................................................ 425
LIST OF APPENDICES

Appendix 2.1 Alignment of RPB2 sequences from Sordariomycete species. Primer-binding sites are indicated by shading................................................................. 65

Appendix 2.2 Alignment of β-tubulin sequences from Sordariomycete species. Priming sites are indicated by shading. ................................................................. 76

Appendix 2.3 RPB2 alignment of *M. nivale* and *M. majus* sequences with HindIII digest sites indicated by shading. ................................................................. 91

Appendix 2.4 Alignment of β-tubulin sequences from *M. nivale* and *M. majus*. Primer-binding sites are indicated by shading. ................................................................. 96

Appendix 2.5 Alignment of EF-1α sequences from *M. nivale* and *M. majus*. Primer binding sites are indicated by shading................................................................. 100

Appendix 2.6 Alignment of ITS sequences from *M. nivale* and *M. majus*. Primer binding sites and *Rsa*I restriction sites indicated by shading. ................................................................. 103

Appendix 3.1 Sample script to execute SOAPdenovo................................................................. 166

Appendix 3.2 Sample configuration file for SOAPdenovo ................................................................. 167

Appendix 3.3 Sample script used to execute ABySS ................................................................. 168

Appendix 3.4 Sample script used to execute Velvet................................................................. 169

Appendix 3.5 Sample script used to execute AUGUSTUS ................................................................. 170

Appendix 3.6 Annotate_genes.pl ................................................................. 171

Appendix 3.7 parse_m9.pl ................................................................. 173

Appendix 3.8 make_simple_table_v2.pl ................................................................. 174

Appendix 3.9 summarize_with_files_v2.pl ................................................................. 177

Appendix 3.10 find_genes_of_diff_length.pl ................................................................. 189
Appendix 3.11 compare_phi_results.pl .............................................................. 191
Appendix 3.12 check_proximity.pl................................................................. 196
Appendix 3.13 eliminate_duplicates.pl.......................................................... 198
Appendix 3.14 Alignment of putative *M. bolleyi* 07020 EF-1α sequence with the *Microdochium* EF-1α primers EFNivF and EFMajF and the reverse complement of primer EFMicR .......... 199
Appendix 3.15 Alignment of predicted gene sequences that are putatively unique to *Microdochium* spp. ........................................................................................................ 201
Appendix 4.1 Sample input for linkage disequilibrium calculation with Arlequin.............. 268
Appendix 4.2 Banding patterns detected with primer groups studied (n= 136). Multiple rows with the same number of bands represent different banding patterns based on band sizes....... 273
Appendix 5.1 Alignment of MAT1-1-1 sequences collected from GenBank for primer design 344
Appendix 5.2 Alignment of MAT1-2-1 sequences collected from GenBank for primer design. ........................................................................................................................................ 353
Appendix 5.3 Alignment of SLA2 coding sequences. Note that trailing sequence has been truncated...................................................................................................................................... 355
Appendix 5.4 Alignment of MAT1-2-1 coding sequences............................................. 362
Appendix 5.5 Alignment of APN2 coding sequences .................................................. 366
Appendix 5.6 Alignment of *M. majus* 99049 MAT-region coding sequences with those of *M. nivale* 11037, with primer loci indicated for SLA2 ................................................................. 371
Appendix 5.7 Alignment of *M. majus* 99049 MAT-region coding sequences with those of *M. nivale* 11037, with primer loci indicated for MAT1-2-1 ......................................................... 376
Appendix 5.8 Alignment of *M. majus* 99049 MAT-region coding sequences with those of *M. nivale* 11037, with primer loci indicated for APN2.............................................................. 380

xxiv
LIST OF ABBREVIATIONS AND ACRONYMS

AFLP: Amplified Fragment Length Polymorphism
BLASTn: Basic Local Alignment Search Tool (nucleotide vs. nucleotides)
BLASTp: Basic Local Alignment Search Tool (protein vs. proteins)
BLASTx: Basic Local Alignment Search Tool (translated nucleotide vs. proteins)
bp: Base Pair(s)
d: day
ddNTP: DiDeoxyNucleotide TriPhosphate
DNA: DeoxyriboNucleic Acid
dNTP: DeoxyNucleotide TriPhosphate
dpi: Days Post-Inoculation
EDTA: EthyleneDiamineTetraacetic Acid
EF-1α: Elongation Factor-1α
FHB: Fusarium Head Blight
IA: Index of Association
ISSR: Inter Simple Sequence Repeat
ITS: Internal Transcribed Spacer
JGI: Joint Genome Institute
h: Hours
HMG: High-Motility Group
hpi: Hours Post-Inoculation
HW: Hardy-Weinberg
KB: Kentucky Bluegrass (*Poa pratensis*)
LSD: Lab Services Division
LSU: Large Subunit
MAT: MAting Type
Mb: MegaBase(s) (i.e. $10^6$ base pairs)
min: Minute(s)
ML: Maximum Likelihood
MP: Maximum Parsimony
mRNA: Messenger RiboNucelic Acid
N50: the contig size at which 50% of all bases in the assembly are contained in contigs / scaffolds that are larger or smaller than this value
NCBI: National Center for Biotechnology Information
NIH: National Institutes of Health
NJ: Neighbour-Joining
NGS: Next-Generation Sequencing
PCR: Polymerase Chain Reaction
Pfam: putataive Protein FAMily
PHI: Pathogen-Host Interaction
PDA: Potato Dextrose Agar
rDNA: Ribosomal DeoxyriboNucleic Acid
rRNA: Ribosomal RiboNucleic Acid
RIP: Repeat-Induced Point mutation
RNA: Ribonucleic Acid
RPB2: Ribonucleic acid Polymerase binding Protein, 2<sup>nd</sup> largest subunit
rpm: Revolutions Per Minute
s: Second(s)
SA: Single-Ascospore
SBS: Sequencing By Synthesis
SDS: Sodium Dodecyl Sulfate
SNP: Single-Nucleotide Polymorphism
SSU: Small SubUnit
TBE: Tris Borate EDTA
tBLASTn: Basic Local Alignment Search Tool (protein vs. translated nucleotides)
tBLASTx: Basic Local Alignment Search Tool (translated nucleotides vs. translated nucleotides)
TE: Tris EDTA
TE: Transposable Element
Tris: tris(hydroxymethyl)aminomethane
× g: times gravity (relative centrifugal force)
UK: United Kingdom
UPGMA: Unweighted Pair Group Method with Arithmetic mean
USA: United States of America
UV: Ultra Violet
v/v: volume of solute per volume of solvent (percent)
VCG: Vegetative Compatibility Group
W: Wheat (<i>Triticum aestivum</i>)
w/v: weight of solute per volume of solvent (percent)
Chapter 1 General Introduction & Literature Review

1.1 Introduction

Plants are necessary not only as food for humans, wild animals, insects, and livestock, but they also perform critical ecological functions, provide materials for construction and textiles, and are valuable sources for both traditional and allopathic medicines. In addition to these vital roles, plants are part of our recreational environments such as sports fields, gardens, and forests. Plants, like all other living things, are subject to disease. Plant pathology is the multidisciplinary field that encompasses the diseases of plants, their causes, mechanisms, and treatments (Agrios 2005). A plant is generally understood to be diseased when its ability to grow or reproduce normally is inhibited by any of a number of different factors, including abiotic threats such as an imbalance in moisture or nutrient levels or by biotic factors including viruses, bacteria, and fungi. Of these threats, fungi are the most common cause of plant diseases (Agrios 2005).

The first fungicide was Bordeaux mixture, a mixture of copper sulfate and calcium hydroxide, which in 1885 was found to suppress powdery mildew on grapes (Agrios 2005). Since this time, a variety of natural and synthetic fungicides have been discovered. However, insufficient understanding of the dangers of some of the chemicals that were used as early fungicides, especially mercurial compounds (Agrios 2005), has contributed to a negative perception of fungicides among the general public (Gullino and Kuijpers 1994; Ragsdale and Sisler 1994). In 2008, the Ontario government banned the use of fungicides and other pesticides for cosmetic uses (Anonymous 2008), although golf courses were exempt from this ban. These changes, in addition to restrictions on individual fungicides (Anonymous 2010), suggest that fungicide use in the future will become more restricted. To facilitate the development of
pathogen-targeted controls, further research into the basic biology of plant pathogens is necessary.

The intensity of fungicide use on golf courses is among the highest in any agricultural or horticultural sector based on the rate per hectare (Anonymous 1998), and by area, more money is spent for turfgrass disease control than for any other cultivated plant (Nelson 1992). In Canada, the greatest single use of fungicides is for snow mould control across the country (Hsiang et al. 1999), and snow moulds are the most economically important turfgrass diseases (Jung et al. 2007). Among the snow mould diseases, grey snow mould caused by *Typhula* species, and pink snow mould caused by *Microdochium* species are the most important (Jung et al. 2007). Pink snow mould and other diseases caused by *Microdochium* on turfgrasses are likely the most common turfgrass diseases in cool, wet climates (Hsiang 2009).

1.2 General information about *Microdochium nivale* and *M. majus*

*Microdochium nivale* (Fr.) Samuels & Hallett and *M. majus* (Wollenw.) Glynn & S.G. Edwards⁠¹ are ascomycete fungal plant pathogens found worldwide in cool and temperate regions (Kammoun et al. 2009; Lees et al. 1995; Nakajima and Naito 1995; Waalwijk et al. 2003). Both species attack cereals (e.g., *Hordeum vulgare* L., *Triticum* spp.), and *M. nivale* is also a pathogen of cool-season grasses (e.g., *Agrostis*, *Lolium*, and *Poa*) (Simpson et al. 2000). Whereas most plant pathogens in temperate regions are not active during the winter (Agrios 2005), *M. nivale* and *M. majus* are active throughout the year, requiring fungicide management strategies (Maloney and Chalupa 2003; Church et al. 2004).

---

¹ These species were previously known as *Fusarium nivale* (Fr.) Sorauer (or *F. nivale* Ces. ex Berl. & Voglino) and *F. nivale* var. *majus* Wollenw., or *Microdochium nivale* var. *nivale* and *M. nivale* var. *majus* (Wollenw.) Samuels & I.C. Hallett, respectively (Glynn et al. 2005).
and *M. majus* continue to grow at temperatures as low as -5 °C (Snider et al. 2000). This adaptation allows these fungi to attack their host plants under snow cover, and pink snow mould is caused by these pathogens.

Pink snow mould is a common disease on graminaceous plants with live aerial parts that are covered by snow for at least six weeks, and is characterized by bleached to orange-brown patches of matted leaf tissue up to 20 cm in diameter which may also display white or pink mycelium (Figure 1.1) (Hsiang 2009). Following wet and cool weather (between 0-15 °C), *M. nivale* may also cause a disease on grasses known as Fusarium patch, also called Microdochium patch. This disease is characterized by irregular bleached or brown-orange patches up to 5 cm in diameter, the edges of which may be bronze in colour when the pathogen is actively growing (Hsiang 2009). Individual patches may coalesce, resulting in large areas of damaged turf. This disease is a serious problem on managed turf surfaces, such as golf courses (Smiley et al. 2005).

In addition to pink snow mould and Fusarium patch, *M. nivale* and *M. majus* are also among the causative agents of Fusarium head blight (FHB), a common and serious disease of wheat and barley (Ioos et al. 2004). Although FHB can be caused by many different pathogens (Pirgozliev et al. 2003), notably *Fusarium culmorum* and *F. graminearum*, *M. nivale sensu lato* is among the most common and severe causal agents (Ioos et al. 2004). European studies have suggested that the incidence of *M. nivale sensu lato* in disease outbreaks has increased in recent years (Ioos et al. 2004). Cereals afflicted with FHB turn brown and wither, and infected kernels are lower in biomass and nutritive value relative to unaffected crops. In severe outbreaks, crop yields may be reduced by 50% (Windels 1999).

The prevention of snow mould diseases, including pink snow mould, accounts for almost 50% of the yearly fungicide use on turfgrass in Canada (Hsiang et al. 1999). Unfortunately,
Despite such management efforts, *M. nivale* and *M. majus* remain common and serious pathogens on wheat and turf in Canada and Europe. Recent studies have demonstrated resistance among *M. nivale* to the strobilurin fungicides that were effective in the past (Walker et al. 2009).

### 1.2.1 Disease Cycle

*Microdochium nivale* and *M. majus* are most likely spread by the dispersal of infected plant materials, including seeds and plant debris such as stem cuttings (Snyder and Nash 1968) which may be left behind following harvest or maintenance. These pathogens persist and grow in soil, and mycelia in particular are known to cause rapid and severe infection relative to conidia (Pronczuk and Messyasz 1991). Ascospores may also constitute an important source of primary inoculum (Mahuku et al. 1998). Infected seeds (Cristani 1992; Humphreys et al. 1995, 1998) may also spread disease from one field to another, although seeds are now regularly tested for fungal infestation and treated with fungicide to mitigate this problem (Glynn et al. 2008). While many fungal plant pathogens use specialized structures or thick-walled spores to survive the winter (e.g., chlamydospores, sclerotia, or ascospores) (Agrios 2005), *Microdochium nivale sensu lato* is a psychrophilic organism capable of growth between approximately -5 and 30°C (Snider et al. 2000), producing mycelium and actively damaging plant tissues under the snow and during the early spring and late fall in temperate climates (Hoshino et al. 2009).

To thrive at cool temperatures, *M. nivale sensu lato* is known to modify the composition of its lipid membrane by incorporating additional triacylglycerols containing unsaturated fatty acids, especially linolenic (18:3) and linolenic (18:2) acids when subjected to low temperatures (Istokovics et al. 1998). It also possesses an unusual betaine lipid that is also found in the snow-mould-causing fungi *Typhula ishikariensis* and *T. incarnata*, although the specific function of
this lipid is unknown (Istokovics et al. 1998). Unlike other cold-tolerant organisms, such as some species of bacteria and amphibians, *Microdochium nivale sensu lato* is not known to produce antifreeze proteins to prevent the formation of ice crystals at temperatures below 0°C (Snider et al. 2000).

### 1.2.2 Phylogenetic classification of *M. nivale* and *M. majus*

*Microdochium nivale* was first identified by Fries (1825), and was given the name *Lanosa nivalis*. This organism was reclassified and placed in the genus *Fusarium* as *F. nivale* (Fr.) Sorauer (1901); however the same name had been used based on a separate type specimen, *Fusarium nivale* Ces. ex Berlese & Voglino (1886). Wollenweber (1930) differentiated a new variety, *F. nivale var. majus*, from the type variety *F. nivale var. nivale*. The lack of a conidial foot cell, a key feature in the genus *Fusarium*, was used as an important character to support the reclassification of the fungus as *Gerlachia nivale* (Gams and Müller 1980). This classification was contested by Samuels and Hallett, who synonymised the genus *Gerlachia* W. Gams & E. Müller with the older genus *Microdochium* Syd. & P. Syd. (1983). A recent examination of the elongation factor 1-α gene, in addition to the morphological and pathogenic differences described elsewhere in the literature (section 1.2), led to the reclassification of *Microdochium nivale var. nivale* and *M. nivale var. majus* as separate species. The former retained the name *M. nivale*, while the latter was named *M. majus* (Glynn et al. 2005).

Under the now-outdated dual nomenclature system for fungi (Taylor 2011), the sexual states (teleomorphs) of *M. nivale sensu lato* have had a similar history of reclassification. The sexual stage of *Fusarium nivale* was named *Calonectria nivalis* Schaffnit in 1913. This name is a synonym of the older name *Nectriella graminicola* (Berk. & Broome) Niessl. The teleomorph
was then transferred to the genus *Griphosphaeria* by Muller & von Arx (1955) due to its production of subepidermal, darkly-pigmented perithecia and to *Micronectriella* by Booth (1971), despite the presence of unitunicate asci in *M. nivalis* (Samuels and Hallett 1983), and bitunicate asci in the type species for *Micronectriella* (Shoemaker 1981). The most recently accepted teleomorph names for *M. nivale* and *M. majus* were *Monographella nivalis* (Mueller 1977) and *Monographella nivalis var. neglecta* (Gams and Müller 1980). The recent "one name - one fungus" resolution of the International Biological Congress has eliminated the dual nomenclature system for fungi (Taylor 2011). Although it is possible that the names of *M. nivale* and *M. majus* may change as a result of this policy, no decision has yet been reached for these species (K. Seifert, personal communication; T. Gräfenhan, personal communication). For this reason, I use the more familiar names *Microdochium nivale* and *M. majus* that traditionally have been associated with the asexual stage (anamorph).

1.3 Differences between *M. majus* and *M. nivale*

Although *M. nivale* and *M. majus* were regarded as conspecific until 2005 (Glynn et al. 2005), consistent differences between these organisms have been known for over 80 years (Wollenweber 1930). Differences exist between *M. majus* and *M. nivale* with respect to host preferences (Simpson et al. 2000), morphology (Litschko and Burpee 1987), production of sexual spores (Smith 1983), and genetic information (Glynn et al. 2005; Maurin et al. 1995; Nicholson et al. 1996; Parry et al. 1995).
1.3.1 Morphological characteristics

On potato dextrose agar (PDA), the hyphae of *Microdochium majus* and *M. nivale* are white to salmon-pink in colour, with a growth rate between 0.13-0.37 mm/h at 20 °C (Litschko and Burpee 1987). Although initially considered to be different varieties of *M. nivale sensu lato*, *M. majus* and *M. nivale* have been recognized as distinct since 1930 based primarily on the difference between their average conidial size (Wollenweber). The conidia of *M. majus* are larger in size (width from 4.2-6.0 µm, length 15-33 µm) than those of *M. nivale* (width no larger than 3.8 µm, length 8-27 µm), and possess more septa (1-7 compared to 0-3) (Gerlach 1982). However, this wide range of characters overlaps, and some researchers have found that many individual isolates of these species fall within an ambiguous range, rendering morphological differentiation alone unreliable (Lees et al. 1995; Litschko and Burpee 1987).

The production of ascospores has been infrequently observed, but some differences may exist between conditions favouring perithecial formation in these two species, where *M. majus* appears to develop perithecia more readily than *M. nivale* (Litschko and Burpee 1987). Nevertheless, conidial variation, despite overlapping attributes, remains the only non-molecular physical characteristic consistently available for distinguishing between these species.

1.3.2 Pathogenic differences

Although both *M. nivale* and *M. majus* are known pathogens of graminaceous plants, consistent differences have been observed in the distribution of these pathogens among their hosts. When inoculated alone or competitively on wheat, oat, and rye seedlings, both *M. nivale* and *M. majus* were capable of causing disease on each host; however, in the mixed inoculation experiments *M. nivale* strongly outcompeted *M. majus* on rye while *M. majus* preferentially
colonized wheat (Simpson et al. 2000). Similarly, in a detached-leaf assay, *M. majus* displayed a faster colonization rate and induced larger lesions than *M. nivale* (Diamond and Cooke 1997). When evaluating the extent of fungal colonization on three cultivars of *Lolium perenne*, *M. nivale* caused more serious infection than *M. majus* (Hofgaard et al. 2006). Populations of *M. nivale* living on different hosts may represent specialized individuals with little gene flow between populations; indeed, *M. nivale* is known to possess a wider host range and exhibits a higher level of genetic diversity relative to *M. majus* (Mahuku et al. 1998). The relationship between *M. nivale*’s genetic diversity and its wide host range has not been well-explored.

**1.3.3 Genetic differences**

Sequence differences between *M. nivale* and *M. majus* have been identified in a number of genetic regions. Specific RFLP profiles of the ITS region (Lees et al. 1995) have been described for *M. nivale* and *M. majus*. Sets of primers designed to amplify only one variety have been developed in an unspecified genetic region (Nicholson et al. 1996) and in the gene for elongation factor 1α (EF-1α) (Glynn et al. 2005). In addition to functioning as a useful tool for the identification of new isolates from the field, these studies have demonstrated the magnitude of variation between *M. nivale* and *M. majus*. Some genetic differences connected to host origin have also been noted. In a study of *M. nivale* isolates from turfgrass, fungal isolates from different host species were more similar to isolates which shared a host than those which were found in close proximity (Mahuku et al. 1998). Conversely, a primer set designed by European researchers to differentiate between *M. nivale* and *M. majus* was unable to amplify DNA from North American *M. nivale* isolates (Glynn et al. 2005). This observation, in combination with the
other differences described above, lead Glynn et al. (2005) to propose that the two varieties *M. nivale* var. *nivale* and *M. nivale* var. *majus* represent distinct species.

1.4 Sexual Reproduction

1.4.1 Sexual reproduction in the Ascomycota

Fungi may be capable of both sexual and asexual reproduction. Asexual propagation may occur in several different ways: via the production of mitospores such as conidia; through the dispersal of hyphal fragments; or through the production of resting structures, such as sclerotia. Sexual reproduction in fungi involves the production of meiospores with novel genetic information (Taylor et al. 1999). Asexual reproduction is common in fungi (Fisher 2007), and some species are only known to proliferate in this manner. Such species were previously classified in the Deuteromycetes since they could not be placed with their closest relatives in earlier classification systems that were based on the morphology of sexual reproductive structures (Schoch et al. 2009). The Deuteromycetes are now known to be a polyphyletic group; this taxon has been abandoned and its members are gradually being placed with their true relatives in a system based on evolutionary relationships (Taylor, 2011).

In contrast to the sexes understood to exist in animals, fungi exhibit a complex mating system that may be controlled by several genes at multiple loci within an individual's genome. These genes are referred to as mating type genes (MAT genes). Among the Ascomycota, there are two different mating types. These genes have been referred to by several different names, including α and a (especially in yeast), + and −, or MAT1-1 and MAT1-2 (Casselton 2008). While both mating type genes may occupy the same genetic region, they are referred to as idiomorphs rather than alleles to emphasize their extremely divergent nature. The mating type regions appear to
code for transcription factors, which in turn regulate the transcription of genes pertaining to the synthesis of hormones, proteins and compounds directly relevant to sexual reproduction (Pöggeler 2000). The two mating type idiomorphs may be identified by their DNA-binding motifs: generally, the MAT1-1 gene contains an α-box motif, while the MAT1-2 gene contains an HMG box (Arie et al. 1999). This observation facilitates the discovery of mating type genes in new species, because the remainder of the gene is often highly divergent (Taylor et al. 1999).

There are three broad categories of sexual reproduction employed by fungi: heterothallism, homothallism, and pseudohomothallism (Butler 2007). In heterothallic fungi, any one individual possesses only a single MAT idiomorph (either MAT1-1 or MAT1-2), and successful sexual reproduction requires two individuals of opposite mating types. Homothallic individuals are self-fertile, and possess both MAT1-1 and MAT1-2 within the same nucleus (although the genes themselves are not necessarily adjacent or in close proximity to each other). Pseudohomothallic individuals possess both MAT1-1 and MAT1-2, and are generally self-fertile, producing ascospores that possess both MAT genes (Merino et al. 1996). However, these genes are carried within separate nuclei. Ascospores that develop into self-fertile individuals receive both nuclei, but, infrequently, ascospores containing only a single nucleus are produced. In this manner, some self-sterile individuals are produced. These self-sterile individuals require an individual of the opposite mating type for successful sexual reproduction, as in heterothallic species (Merino et al. 1996).

Modern genetic techniques have facilitated the identification of cryptic sexual reproduction in populations previously believed to be strictly asexual (Arie et al. 2000). One expected quality of a truly asexual population would be relatively low genetic diversity, as new genetic information would be expected to arise from mutation alone, rather than via meiotic
recombination as in sexual populations (Taylor et al. 1999), although a large enough asexually reproducing population may possess a similar level of genetic diversity relative to a sexual population (e.g., (Groth et al. 1995).

**1.4.2 Sexual reproduction in Microdochium nivale and M. majus**

The relatively large genetic variability observed in *M. nivale* is consistent with a sexual population, but sexual reproduction has rarely been observed in this species (Booth 1971; Cook and Bruehl 1966; Mahuku et al. 1998). Both *M. nivale* and *M. majus* produce perithecia in the lab (Lees et al. 1995; Litschko and Burpee 1987; Parry et al. 1995); however, *M. majus* appears to produce perithecia more readily than *M. nivale*, as it appears to be homothallic (self-fertile). Old literature, which may have been referring to *M. majus*, describes *M. nivale* as homothallic; however, more recent research has suggested that *M. nivale sensu stricto* may actually be heterothallic, requiring individuals of opposite mating types for successful reproduction (Lees et al. 1995). If true, this could help to explain the relatively large degree of heterogeneity observed in *M. nivale* relative to *M. majus*. Furthermore, as both species actively cause disease at low temperatures, ascospores may serve as survival spores during warm, dry conditions which do not favour infection by these pathogens.

If *M. majus* is truly homothallic while *M. nivale* is heterothallic, then both mating type idiomorphs should be found in the genome of a single *M. majus* individual while only a single idiomorph would be expected in a *M. nivale* individual. The identification of mating type genes in these species has not yet been reported in the literature, but could be useful in understanding the degree of genetic diversity present in this pathogen. This information would be timely as resistance to the strobilurin pesticides has been recently observed (Walker et al. 2009);
understanding the mechanism of how the genes responsible for resistance may spread through the population could help to mitigate the transmission of this potentially harmful mutation.

1.5 Sequencing techniques and bioinformatics

1.5.1 DNA sequencing techniques

Beginning with the first viral genome sequenced in 1977 (Sanger et al. 1977a), whole-genome sequencing has allowed researchers to better understand individual species and the relationships between them. Although several different sequencing methods were explored by researchers in the 1970s, the automation of the Sanger method allowed this technique to become the method of choice for most sequencing applications until the early 21st century (Shendure and Ji 2008). In Sanger sequencing, the sequence of interest is amplified in a buffered solution containing a mixture of standard deoxynucleotides (dNTPs) and chain-terminating dideoxynucleotides (ddNTPs) in the presence of a DNA polymerase enzyme and a primer that is specific to the sequence of interest.

When the method was first developed, the ddNTPs were labelled radioactively, and separate reactions were conducted for each of the four nucleotides (Sanger et al. 1977b). Within each reaction mixture, as the polymerase synthesizes a copy of the sequence of interest, the labelled ddNTP will be randomly incorporated into the sequence in the place of the correct dNTP. As this process is repeated on the millions of template DNA molecules included in the reaction mixture, the amplification products will consist of truncated sequences that vary in length based upon when (or if) the labelled nucleotide was incorporated into the growing product. This mixture of sequences can then be separated according to size using gel electrophoresis and, in combination with the results from the other labelled nucleotides, the sequence of the original template strand
may be determined. Modern methods use ddNTPs labelled with four different fluorescent dyes, which eliminates the need to run a separate reaction for each nucleotide and also facilitates automatic detection (Shendure and Ji 2008).

The Sanger method has been used to generate full genome sequences for an assortment of diverse species, including humans. Although Sanger sequencing remains in wide use, a number of alternative sequencing strategies have become available in the past 10 years, and these techniques offer several improvements over the Sanger sequencing method in terms of both the cost and the length of time required to produce data (see section 3.1) (Lister et al. 2009). These newer technologies are broadly referred to as "next-generation" sequencing (NGS) (Shendure and Ji 2008). These techniques are more prone to certain types of sequencing error than Sanger sequencing, and the sequence fragments produced are generally shorter than those produced by the Sanger method, but they can be produced much less expensively and allow for much greater coverage (Shendure and Ji 2008). Advances in data processing methods (see section 1.5.3) have allowed sequence reads from these newer techniques to be used in de novo genome assembly to produce complete genome sequences for eukaryotes, including fungi (Nowrousian et al. 2010), plants (Shulaev et al. 2011), and mammals (Li et al. 2009).

Filamentous ascomycetes are particularly well-suited for whole-genome sequencing because of their monokaryotic haploid vegetative growth (Webster and Weber 2007), which facilitates the collection of a large amount of genetically homogenous material for sequencing. In addition, fungal genomes are typically less than 100 Mb in size (Haridas et al. 2011), which decreases both the complexity of the assembly process and the cost and resources required to obtain sufficient information for successful assembly.
1.5.2 RNA sequencing

While the analysis of DNA sequences facilitates the analysis of relationships between different species or individuals, the identity and the abundance of RNA transcripts describes the priorities of an individual organism under a specific set of conditions (Wang et al. 2009). Within the context of a plant-pathogen interaction, both the host plant and the pathogen undergo changes in protein synthesis when the host is challenged by a pathogen. For example, the host may synthesize defensive or reparative enzymes while the pathogen may secrete cellulases and other enzymes responsible for degrading the host plant (Conrath et al. 2002; Dickinson 2003; Jones and Dangl 2006). The transcriptome of an organism is defined as the totality of its RNA pool under a specific set of conditions, including both the identities and the quantities of all transcripts (Wang et al. 2009). By comparing the transcriptome of healthy to that of diseased host plant cells, it may be possible to uncover genes which play a role in the plant’s defensive response. In addition to identifying differences in the identities of the mRNA transcripts present, it is also possible to identify differences in the relative quantities of each transcript (Li et al. 2010). If an organism is examined at different times, or in response to different stimuli, it may be possible to identify genes that have been up- or down-regulated in response to these changes.

Transcriptome analysis generates sequence information for all of the mRNA transcripts present in the organism at the time of sampling (Grabherr et al. 2011), within the limits of detection of the technique used. This powerful technique has been made possible through the recent NGS developments discussed above, and offers many advantages over previously-developed transcriptome analysis techniques such as northern blotting, microarrays, and quantitative reverse transcriptase PCR. These older techniques required prior knowledge of the sequences of interest, or of specific probes or primers, or were limited by their range of detection
(Wang et al. 2009). These restrictions made it difficult to detect transcripts which may be infrequent or unique, whereas NGS techniques excel at very high levels of coverage which might reveal such transcripts especially through the use of in silico subtraction techniques.

Transcriptome sequencing using NGS techniques, known as RNA-Seq (Nagalakshimi et al. 2008), necessitates the extraction of the RNA pool of an organism. The mRNA is sequestered by affinity chromatography targeting the poly-A tail that is added in post-transcriptional modification (Nelson and Cox 2004). The mRNA is then used as a template to synthesize cDNA using reverse transcriptase, and NGS techniques are used to sequence the entire cDNA pool.

The power of RNA-Seq has been demonstrated by generating sequence libraries for yeast both under vegetative (Nagalakshimi et al. 2008) and meiotic conditions (Wilhelm et al. 2008). Comparisons between the transcriptomes of wild-type versus perturbed organisms have been made by many researchers using different experimental conditions. For example, the small RNA (smRNA) complements of cold-tolerant and cold-susceptible Triticum aestivum challenged with cold temperatures were investigated, and researchers found significant differences in transcript levels between the cultivar types (Qin et al. 2008). This validation of transcriptome analysis suggests that further studies could reveal important molecular processes involved in disease and stress tolerance. Importantly, both Microdochium majus and M. nivale are known wheat pathogens, but when grown competitively, M. majus out-competes M. nivale (Simpson et al. 2000): this distinction is likely linked to differences within the pathogens themselves as well as in the response mounted by the host. Transcriptome analysis of wheat that has been inoculated with different pathogens may reveal differences in host response which allow M. majus to colonize more successfully; additionally, by identifying genes and processes important in the
wheat’s defensive response, it may be possible to develop effective breeding programs to select for cultivars that overexpress those genes important for resistance.

1.6 Hypotheses and Objectives

Hypotheses

1. An examination of multiple genes will support the reclassification of *M. nivale* and *M. majus* as distinct species. Furthermore, an examination of the genomes of these organisms will reveal further differences between these species. Populations of *M. nivale* and *M. majus* in North America are genetically distinct from those found in Europe, and *M. nivale* isolates found on turfgrass are genetically distinct from those found on wheat. Repeated sampling of *M. nivale* populations over several years will reveal year-to-year genetic variability at a level that is consistent with a sexually reproducing population.

2. Mating-type genes for *M. majus* and *M. nivale* will be revealed by genome sequencing, and the alternate idiomorph can be uncovered by sequencing the genes surrounding the mating type region in other isolates.

3. Differences in host preferences reported for *M. nivale* and *M. majus* will be reflected in differing infection processes or timing when grown on Kentucky bluegrass (*Poa pratensis*) or on wheat (*Triticum aestivum*).

Objectives

1a. Obtain isolates of *M. nivale* and *M. majus* from around the world and from different hosts (both wheat and turfgrasses). Amplify and sequence a variety of protein-coding and non-protein-coding genes. Analyze the sequence differences between *M. majus* and *M. nivale,*
and compare these differences to those observed for an outgroup species. Use this information to assess whether *M. nivale* and *M. majus* possess sufficient heterogeneity to justify their elevation to species, and to assess the presence or absence of sequence divergence among different host and geographic populations. (Chapter 2)

1b. Obtain complete genome sequences for one isolate of *M. nivale* from wheat, one isolate of *M. nivale* from turfgrass, and one isolate of *M. majus* from wheat using NGS, and identify divergent regions between the genomes of *M. nivale* and *M. majus* to further assess the elevation to full species and to search for genes that may be unique to these two species, individually and collectively. (Chapter 3)

1c. Assess genotype performance by examining genetic diversity among and between populations sampled over multiple years using ISSR markers. (Chapter 4)

2. Identify the mating-type genes within the *M. nivale* and *M. majus* genomes based on sequences obtained from other filamentous ascomycetes. Sequence and compare mating types from individual isolates of *M. nivale* and *M. majus*, and screen the in-lab isolate collections of these species to classify isolates according to mating type. Attempt *in vitro* crosses between individuals of opposite mating types (if available) and study mating process microscopically if mating can be induced. (Chapters 3 and 5)

3. Inoculate detached leaves of Kentucky bluegrass and wheat with hyphal suspensions of *M. nivale* and *M. majus* to study the infection process. Collect leaf samples and examine microscopically to determine the timing and mechanism of infection for these pathogens on two different host plants. (Chapter 6)
1.7 References for Chapter 1

Anonymous. 2008. An Act to amend the Pesticides Act to prohibit the use and sale of pesticides that may be used for cosmetic purposes. Edited by Province of Ontario.
Cristani, C. 1992. Seed-borne Microdochium nivale (Ces. ex Sacc.) Samuels (= Fusarium nivale (Fr.) Ces.) in naturally infected seeds of wheat and Triticale in Italy. Seed Science and Technology 20(3): 603-617.


Hsiang, T. 2009. All you ever wanted to know about Fusarium patch / Microdochium patch / pink snow mold or whatever that disease is called. Green Master **44**(4): 13-16.


Pronczuk, M., and Messyasz, M. 1991. Infection ability of mycelium and spores of Microdochium nivale (Fr.) Samuels & Hallett to Lolium perenne L. Mycotoxin Research 7A.


Figure 1.1 Pink snow mould damage on mixed creeping bentgrass / annual bluegrass green (Agrostis stolonifera / Poa annua), Guelph, ON, 02 March 2012.
Chapter 2 Phylogenetic Analyses

2.1 Introduction

2.1.1 Fungal taxonomy and nomenclature

Taxonomy describes the process of creating hierarchical groups of living things, whereas nomenclature is the system that governs the naming of the organisms and groups delineated by taxonomy (Lizon and Samuels 1997). Historically, fungi were identified and classified according to their morphology, especially that of the sexual spores and structures (Webster and Weber 2007). The first set of rules establishing the nomenclature of fungi were published in the late 19th century (de Candolle 1867) in what was to become the International Code of Botanical Nomenclature (ICBN). Under the versions of the ICBN in force from 1905 until 2011, it was permitted to apply separate names to the sexual and asexual stages of the same organism (the teleomorph and anamorph name, respectively), and these different generic names were separately classified at higher ranks as well. Recent proposals for a more sensible one fungus - one name system (Hawksworth 2004; Taylor 2011) were adopted in the International Code of Nomenclature for algae, fungi, and plants (Hawksworth 2004; McNeill et al. 2012), ending the dual nomenclature system that had been employed for fungi for over 100 years.

This reliance on knowledge of sexual structures for classification meant that species which were not known to reproduce sexually could not be classified with their closest relatives. The phylum Deuteromycota was erected to house these asexual species (Webster and Weber 2007). However, the currently-accepted fungal phylogeny does not include the Deuteromycota because these fungi are asexual forms of Ascomycota or Basidiomycota (Hibbett et al. 2007). Although morphological observations are still valuable for the identification of fungi, the
addition of molecular methods (see section 2.1.2) has revolutionized fungal taxonomy (Schoch et al. 2009).

2.1.2 Molecular phylogeny

The use of DNA, rather than morphology, for taxonomic distinctions among the fungi was pioneered in the early 1990s (Hibbett et al. 2007) by White and colleagues (1990). The addition of molecular characters to the morphological and ecological observations that had been previously used to delineate fungi into taxa revealed the presence of cryptic species in several economically and environmentally important groups, including human pathogens (Koufopanou et al. 2001; Pringle et al. 2005), agricultural pests (Wang et al. 1998), and ecologically sensitive species (Bickford et al. 2007).

In recent years, the amount of genetic information available in public databases such as GenBank (run by the National Center for Biotechnology Information; www.ncbi.nlm.nih.gov/) has increased rapidly due to technical improvements and the concurrent decrease in sequencing cost. While this enormous pool of data has been beneficial for researchers in many disciplines, the most effective ways to manage, analyze, and interpret this information have been contentious. For example, when sequence information for multiple genes is available for the taxa of interest, some researchers have questioned whether it is more effective to compile all available data into a single analysis, or if each genetic region should instead be considered separately. If these regions are analysed separately, conclusions about the dataset as a whole may be derived from the results each individual analysis (Gontcharov et al. 2004). The goal of such experiments, often using cladistic methods, is to discover the “true” tree which is in agreement with the evolutionary history of the organisms under study (Schuh and Brower 2009). However, except in
very rare cases, the true evolutionary history is unknown. It is therefore important to carefully consider the assumptions implicit in any analysis in order to ensure that the results are meaningful.

2.1.3 Genes used for molecular phylogeny in fungi

The rDNA region has been of interest for phylogenetic study since the 1980's (Woese and Olsen 1986). This region is highly appealing for taxonomic study for several reasons. First, it is of high copy number in the genome, which reduces the total amount of DNA necessary for successful amplification (Nilsson et al. 2008). The rDNA region consists of highly conserved regions (18s and 28s) interspersed with less-conserved regions (the internal transcribed spacer [ITS] and intergenic spacer [IGS] regions), which facilitates the design of universal primers (White et al. 1990) and provides several levels of taxonomic resolution (Nilsson et al. 2008). Furthermore, at the present time, this region is the most extensively sequenced and studied in fungi (e.g. (Schoch et al. 2009)), providing an enormous pool of data for researchers to use in either fungal identification or for taxonomic studies (O'Brien et al. 2005). For example, the International Barcode of Life project (http://www.barcodeoflife.org/) recently selected the ITS region as a universal marker region for the rapid identification of fungi (Schoch et al. 2012). In recent re-evaluations of the taxonomies of the fungi (Hibbett et al. 2007) and the Ascomycota (Schoch et al. 2009), rDNA sequences were included as important sources of evidence for the proposed re-ordering of several major taxa.

Despite this frequent use, the rDNA region alone, especially the sometimes highly variable ITS region, may not be sufficient for low-level taxonomic distinctions. For example, the 95-97% sequence similarity proposed as the threshold for the distinction between species may
fall within the range of intraspecific variation in some species (Nilsson et al. 2008; Wang et al. 2011). Furthermore, the ITS region of some taxa from different genera may be highly similar (e.g. *Magnaporthe poae* and *Gaeumannomyces graminis*) (data not shown). Supplemental information from ecology or morphology, or from additional genetic regions, is thus valuable in inter- and intraspecific distinction.

In addition to rDNA sequences such as 18S or 28S, protein-coding genes such as β-tubulin (Keeling et al. 2000; Myllys et al. 2002), RNA-polymerase binding proteins 1 and 2 (Hibbett et al. 2007; Schoch et al. 2009), elongation factor-1α (Schoch et al. 2009), glyceraldehyde-3-phosphate dehydrogenase (Myllys et al. 2002) and cytochrome C oxidase 1 (Damon et al. 2010; Seifert et al. 2007), have also been used extensively in molecular phylogeny among the fungi.

### 2.1.4 Tree-building algorithms

The construction of phylogenetic trees provides a visual representation of a hypothesis describing the relationships between various taxa. The choice of algorithm used for the construction of a tree may affect the apparent relationships between the taxa under study. Several tree-building algorithms are currently in use, each with a different balance of strengths and weaknesses.

Neighbour joining (NJ) is an example of a distance matrix-based method, where an algorithm is used to calculate the relative distances between the taxa under study based on differences in their character states (eg. nucleic acid sequence) (Holder and Lewis 2003). In NJ, taxa are combined into a group if their combination decreased the overall tree length. This method is repeated until the relationships between all of the taxa have been computed (Saitou
and Nei 1987). Bootstrap analysis, wherein the dataset is re-sampled many times, is often conducted to estimate the statistical support for the divisions proposed by the tree (Felsenstein 1985). Neighbour-joining is generally a rapid and non-computationally demanding method (Holder and Lewis 2003), but it is insensitive to homoplasy (Holder and Lewis 2003), and the resulting tree may be influenced by the order of the input sequences (Harrison and Langdale 2006; Schuh and Brower 2009).

In maximum likelihood (ML), a likelihood function is defined according to the number of sites present in the sequences. Evolutionary events such as transitions (changing from one pyrimidine or purine into another), transversions (changing from a purine into a pyrimidine or vice versa), and indels (the insertion or removal of a nucleotide), are assigned independent probabilities as they may occur at different rates (Holder and Lewis 2003). The likelihood function is evaluated such that the tree with the largest likelihood value – which minimizes “unlikely” mutations such as indels at the expense of maximising “likely” mutations such as transitions – is the most likely to have produced the observed data, and is chosen as the best tree (Li 1997). Given sufficient time and computational power, ML will find the "true" tree (Li 1997); however, ML is a computationally intensive method and it is not computationally feasible to evaluate all of the possible trees in any but the smallest datasets (Holder and Lewis 2003), which means that ML may fail to find the optimal tree.

In maximum parsimony (MP), the optimum tree is the one that requires the smallest number of evolutionary events to explain sequence differences at parsimony-informative sites (Holder and Lewis 2003). When the sequences of interest have been aligned, parsimony-informative sites are those which possess at least two different nucleotides (or amino acids), each of which are present in at least two, but not all, of the sequences under study (Fitch 1977). The
minimum number of changes (transitions, transversions, or indels) required to explain the nucleotide changes across all of the sequences are recorded for that site. This process is repeated at each phylogenetically informative site in the sequence data to construct a tree (Li 1997). Maximum parsimony is intermediate in computational demand between ML and NJ, but the results that it produces may be biased by long-branch attraction, wherein chance similarities between divergent sequences may result in them being incorrectly placed close together on the MP tree (Holder and Lewis 2003). As in NJ and ML, bootstrapping should be used to estimate the statistical strength of each branch within a MP tree (Holder and Lewis 2003).

Despite their individual weaknesses, all three of the methods described are still in regular use for estimating the relationships between taxa (Hibbett et al. 2007; Schoch et al. 2009). A search of the google scholar database using the term phylog* and either "maximum likelihood", "maximum parsimony" or "'neighbour joining' OR 'neighbor joining'" for articles published between 2000 and 2012 revealed that neighbour joining was the most commonly referenced method until 2009, when it was overtaken by maximum likelihood (Figure 2.1).

2.1.5 Genetic differences between and within M. nivale and M. majus

As there are few characteristic morphological differences between M. nivale and M. majus (Gerlach 1982), a number of molecular markers have been described to study variability within and to differentiate between them. Maurin (1995) studied a diverse range of isolates from across Europe and from a number of plant hosts using a range of physical and biochemical characteristics including esterase enzyme profiles and RFLP profiles of the ITS rDNA. Amplification of the ITS region followed by digestion with the restriction enzyme RsaI made a single cut in some of the isolates, including a reference sample of var. majus, while it failed to
cut other isolates, including the reference sample for var. *nivale*. The esterase profiles visualized on acrylamide gels provided another distinguishing feature: while esterase profiles of *M. majus* were generally similar, those of *M. nivale* displayed more heterogeneity. Primers specific to *M. nivale* and to *M. majus* were designed by Nicholson (1996) based on primers used for RAPD PCR. These primers have failed to amplify North American isolates of *M. nivale* (Glynn et al. 2005; Jewell and Hsiang 2013) suggesting possible heterogeneity between the North American and European populations of this species.

To elucidate the genetic diversity and host specificity of *M. nivale* in south-eastern Ontario, numerous isolates of *M. nivale* were examined by Mahuku (1998). Isolates were collected from *Lolium perenne* and *P. pratensis* in Guelph and from *A. stolonifera* (in Guelph and in Cambridge, both in Ontario, 25 km apart). All isolates were identified as *M. nivale* using the *RsaI* restriction assay, and digestion of the IGS region with the restriction enzymes *Hae* III, *Cfo* I, and *Mbo* I identified 60 unique genotypes among the 100 isolates studied. Random amplification of polymorphic DNA (RAPD) PCR revealed the presence of 96 unique genotypes, and dendograms constructed from the RAPD data separated the isolates into four groups, corresponding to the population from which they had been initially collected. Interestingly, the populations from Guelph and Cambridge which shared the same host species were more similar to one another than to any of the other isolate groups from other host species, implying that modest host specialization may be present within turfgrass-colonizing *M. nivale* strains (Mahuku et al. 1998). The large number of unique genotypes observed suggests that sexual reproduction may occur in this species despite limited direct field evidence to support this claim. Furthermore, the large degree of intra-varietal heterogeneity observed in this study agrees with the assertion by Maurin et al. (1995) that *M. nivale* isolates are genetically heterogeneous.
Glynn et al. (2005) examined a portion of the EF-1α coding region that was 838 base pairs in length for 15 isolates of *M. nivale* sensu lato by designing specific primers to amplify either *M. nivale* or *M. majus*. Sequencing of the resulting amplicons revealed 96% sequence similarity between *M. nivale* and *M. majus*, while the mean sequence similarities within each group were 99.7% and 99.8%, respectively. In addition to developing a useful identification tool, this study prompted these researchers to suggest that the subspecies of *M. nivale sensu lato* should be promoted to the species level of classification (Glynn et al. 2005). However, this assertion is based upon the analysis of a single gene; in contrast, a recent revision of general ascomycete taxonomy relied on six genes (Schoch et al. 2009). Literature published on *M. nivale* and *M. majus* since Glynn et al.’s assertion have not always accepted the elevation (eg. (Kaneko and Ishii 2009; Pociecha et al. 2010)). Additional studies using multiple genes to confirm this divergence would support the elevation of *M. nivale* and *M. majus* to distinct species.

### 2.1.6 Objectives

The first major objective of this project was to determine whether an examination of four genetic regions would reveal differences between isolates of *M. majus* and *M. nivale* that are consistent with their elevation to species from varieties. Intra-specific variation was assessed by investigating isolates originating from both Europe and North America and, for *M. nivale*, for isolates originating from turfgrass and from wheat and from both continents. The utility of previously-published primers for distinguishing between *M. nivale* and *M. majus* was assessed.
2.2 Materials and Methods

2.2.1 Sample Collection

Isolates of *M. nivale* and *M. majus* were from the local lab collection (preserved on frozen grain at -20°C), from the field, or from collaborators (Table 2.1). Field isolates were obtained from plants as follows: leaf blades were cut into small pieces no larger than 0.5 cm in length, rinsed in sterile distilled water (SDW), surface sterilized in a 1% NaOCl solution for 0.5–5 minutes, rinsed in fresh SDW, and placed onto 2% potato dextrose agar (PDA, Becton, Dickinson and Company, MD, USA) prepared according to the manufacturer’s instructions and amended with the antibiotics tetracycline (0.5 µg/mL) and streptomycin (1 µg/mL). The plates were stored at 10 °C in the dark for up to two weeks until fungal colonies were visible; specimens resembling *M. nivale* or *M. majus* were isolated by cutting a small plug of agar from the actively growing margin of the colonies and placing them in the centre of a Petri dish containing unamended 2% PDA. Several transfers were made and, where possible, single-spore isolates were obtained before subsequent experiments and for preparing stock tubes for the permanent collection.

Fungal isolates were identified on the basis of growth rate on PDA (0.13-0.37 mm/h at 20 °C) (Litschko and Burpee 1987), colony morphology (salmon-pink mycelium), and, where available, conidial morphology (hyaline, 0-3 septa with width < 4.7 µm, length 8-27 µm for *M. nivale* and hyaline, 1-7 septa with width > 4.5 µm, length 15-33 µm for *M. majus*) (Glynn et al. 2005). All fungal isolates were maintained on 2% PDA and stored at 4 °C. For long-term storage, 15 mL vials containing thrice-autoclaved wheat seeds were inoculated with a fungal isolate, or agar plugs were placed into 15 mL vials containing SDW (5 mL). One vial each of water and wheat-seed were stored at 4 °C while two additional wheat vials were stored at -20 °C.
A total of 62 isolates were included in these experiments (Table 2.1).

2.2.2 DNA extraction

Fungal DNA was obtained by growing each isolate on PDA overlaid with a 6 cm x 6 cm cellophane sheet (Flexel Inc., Atlanta, Ga., U.S.A.) presterilized by autoclaving three times in dH₂O (20 minutes, 121°C). After 5 days, mycelium was harvested by scraping it from the cellophane (avoiding the initial agar plug), and collecting in a 1.5 mL tube. Mycelium was stored at -20 °C until DNA extraction (Edwards et al. 1991). For each DNA extraction, 100 mg of either fresh or frozen mycelium was placed in a 1.5 mL tube containing approximately 50 mg of acid washed and autoclaved sea sand (Fisher, Fair Lawn NJ, USA) and 200 µL Edwards buffer (Tris HCl (pH 7.5) 200 mM, NaCl 250 mM, EDTA 25 mM, SDS 0.5% (w/v)). The mycelium was ground with a plastic pestle (Froggabio, Toronto, Canada) using a Mastercraft Lithium Ion screwdriver (3.6V, Canada) for between 60 to 120 s to disrupt the cells. A further 200 µL of Edwards buffer was added and the extraction mixture was incubated at room temperature for 2 - 3 h, after which it was centrifuged at 12,000 x g in an Eppendorf 5415 D centrifuge (Eppendorf, Mississauga, Canada) for 10 min. The supernatant was transferred into a fresh 1.5 mL tube and an equal volume of -20 °C isopropanol was added to precipitate the DNA. The tube was incubated at -20 °C for at least 1 h. The tube was then spun in the centrifuge at 12,000x g for 10 min to pellet the DNA. The supernatant was discarded and the pellet was washed with 200 µL of cold 70% ethanol. The ethanol was discarded and the pellet was dried by inverting the tube and incubating at room temperature for 10 min. The DNA was then resuspended in 100 µL of PCR water (nuclease free water, Fisher Scientific, USA) by gently pipeting several times. The tubes
were incubated at 4 °C for at least 1 h to allow the dissolution of the DNA before storage at -20 °C.

DNA quality and quantity were assessed by subjecting the samples to electrophoresis through a 1% agarose gel (UltraPure™ Agarose, Invitrogen, Carlsbad, CA, USA) prepared using 0.5X TBE buffer (90 mM Tris base, 90mM boric acid, 2mM EDTA). A 3 μL mixture (2:1) of DNA and loading buffer (0.1% w/v bromophenol blue, 0.1% w/v xylene cyanol, 30% v/v/ glycerol, and 60 mM EDTA) was loaded onto the gel and subjected to electrophoresis in 0.5X TBE buffer at 100 V for approximately 30 minutes before being stained with ethidium bromide and visualized under UV light. The size and approximate concentration of the amplicons were calculated by comparison to a molecular weight and mass ladder loaded in every gel (GeneRuler 100bp DNA ladder, Fermentas, Canada).

2.2.3 Primer design and selection

To examine the genetic regions included in this study, previously published fungal primers were tested in *M. nivale* and *M. majus* (Table 2.2) using the PCR protocol suggested in the original journal article where the primers were designed. In cases where these primers failed to yield consistent results among all of the isolates examined, new primers were designed for each particular region.

PCR primers were designed according to the following general protocol. Sequences from the gene of interest were downloaded from the NCBI GenBank database. When possible, sequences were selected from *Microdochium* species; when no *Microdochium* sequences were available, sequences from other taxonomically related species were chosen. The sequences were aligned (section 2.2.6) and were visualized in BioEdit. Candidate primer sequences were selected
within the conserved regions predicted by the alignment. The candidate primers were assessed using both Primer3 v. 2.2.2 (Rozen and Skaletsky 2000) and OligoCalc (Kibbe 2007). Primers that were between 16-22 bp in length, possessed a GC content between 40-70%, had a predicted melting temperature between 48-64 °C, and which were not predicted to possess self-complementarity or to fold into hairpin structures were selected as strong candidates. Primer pairs that were to be used in a single reaction were further screened to ensure that their predicted melting temperatures were within 2 °C and that they were unlikely to form dimer structures. All primers included in these experiments were ordered from Laboratory Services Division, University of Guelph.

2.2.4 PCR protocols and sequencing

All PCR reactions were optimized to yield single amplicons. The general PCR procedure utilized was as follows: approximately 1 ng (1μL) of fungal DNA was added to a 0.2 mL PCR tube containing a pre-mixed solution consisting of 1X PCR buffer, either 0.2 mM (ITS, EF-1α) or 0.4 mM (β-tubulin, RPB2) Mg^{2+}, 2.5 mM dNTP mixture (prepared using individual solutions of dATP, dTTP, dCTP, and dGTP; BioBasic, Markham, ON), 0.5 μM of each the forward and reverse primers, 0.04 U of Tsg DNA polymerase (BioBasic, Markham, ON), and enough sterile water to bring the total volume (including DNA) to 15 μL. The reaction mixture was subjected to a thermal cycling procedure. For ITS, the thermal cycling protocol consisted of five minutes at 95 °C, followed by 30 cycles of a 30 s denaturation at 95 °C, a 45 s annealing at 55 °C, and a 60 s extension at 72 °C, followed by a final 10 minute extension at 72 °C. For RPB2, the thermal cycling protocol consisted of five minutes at 95 °C, followed by 35 cycles of a 30 s denaturation at 95 °C, a 45 s annealing at 58 °C, and a 90 s extension at 72 °C, followed by a final 10 minute
extension at 72 °C. For β-tubulin, the thermal cycling protocol consisted of five minutes at 95 °C, followed by 35 cycles of a 30 s denaturation at 95 °C, a 45 s annealing at 53 °C, and a 90 s extension at 72 °C, followed by a final 10 minute extension at 72 °C. For EF-1α, the thermal cycling protocol consisted of 75 s at 95 °C, followed by 35 cycles of a 30 s denaturation at 95 °C, a 15 s annealing at 52 °C, and a 45 s extension at 72 °C, followed by a final 4.3 minute extension at 72 °C. For Y13, the thermal cycling protocol consisted of five minutes at 95 °C, followed by 35 cycles of a 30 s denaturation at 95 °C, a 60 s annealing at 45 °C, and a 45 s extension at 72 °C, followed by a final 5 minute extension at 72 °C.

The PCR products were visualized as described above except that PCR products larger than 1kb were subjected to electrophoresis at 50 V for 60 minutes before being stained with an ethidium bromide solution and visualized under UV light. The PCR products were sequenced in the forward direction only by Laboratory Services, University of Guelph (Guelph, Ontario, Canada). Sequencing was performed on a GeneAmp® PCR System 9700 or 2720 Thermal Cycler (Applied Biosystems).

2.2.5 Sequence alignments and trees

Chromatogram files from sequencing reactions were visualized using BioEdit Sequence Alignment Editor software v.7.0.53 (Hall 1999), and base calls were corrected manually as necessary. Within each genomic region studied, the longest sequence common to all of the sequenced amplicons obtained was used for alignment. Alignments were performed using the default parameters with ClustalX v. 2.0.12 (Chenna et al. 2003). Neighbour-Joining (NJ), Maximum Likelihood (ML), and Maximum Parsimony (MP) trees were constructed using PAUP* v. 4.0 b. 10 (Swofford 2000) using a heuristic search with random addition of sequences.
and 1,000 bootstrap replicates. For the maximum likelihood calculations, a transition:transversion ratio of 2:1 was used. For the Neighbour-Joining analysis, the HKY85 model was used. For the RPB2, β-tubulin, and ITS regions, DNA sequences from North American isolates of *Microdochium bolleyi* from our isolate collection were used as outgroups. In the EF-1α analysis, a sequence from *Hypocrea lixii* (DQ056745.1) was used as an outgroup due to difficulty in amplifying the same portion of the EF-1α gene from *M. bolleyi* that was analysed for *M. nivale* and *M. majus*.

### 2.2.6 Restriction Digests

A restriction digest of the RPB2 PCR product was developed using the restriction enzyme *HindIII*. Following amplification with the RPB2 primers RPB150F and fRPB2-7cR, the 1,040 bp amplicon was digested by incubating 8 μL of the PCR product for 2 hours at 32°C in a solution containing 0.02 U *HindIII* restriction enzyme (Invitrogen, Carlesbad, CA), 1x restriction buffer provided by the enzyme's manufacturer, and enough sterile dH₂O to bring the final volume of the solution to 40 μL. Following incubation, the reaction mixture was heated to 72 °C for 10 minutes to inactivate the enzyme. The digested samples were visualized as described for the PCR samples.

### 2.3 Results

#### 2.3.1 Primer testing and design

At the beginning of these experiments, published primer sequences were tested to determine whether the regions of interest could be consistently amplified from a range of isolates. For both ITS and EF-1α, the published primers produced amplicons of the predicted
sizes without optimization. For RPB2, the literature primer set fRPB2-7cF / fRPB2-7cR was tested with 15 isolates of *M. nivale* and 3 isolates of *M. majus*. Following optimization of the PCR protocol, two of the three *M. majus* isolates tested were successfully amplified, but only 3 of the 15 *M. nivale* isolates tested produced a band of the predicted size (1,500 bp). The remaining 12 *M. nivale* and the other *M. majus* isolate failed to yield any amplicon.

The RPB2 amplicon from one of the *M. majus* isolates was submitted for sequencing in the forward direction (LSD) and a 917 bp sequence was returned. When searched against the GenBank database, the top match was the RPB2 sequence from *Pseudomassaria carolinensis* (Order Xylariales, Family Nectriaceae; GenBank accession DQ810239.1). A total of 20 RPB2 sequences were collected from GenBank (Table 2.3) and were aligned with the putative RPB2 sequence from *M. majus* to design a new forward primer. Following optimization, this primer, when paired with fRPB2-7cR, yielded a single band of the predicted size from all of the *M. nivale* and *M. majus* isolates tested.

Similarly, the β-tubulin primers Bt1b / Bt2a (Glass 1995) were tested with a total of five *M. majus* and six *M. nivale* isolates. Following optimization of the PCR protocol, among the *M. majus* isolates, two out of five were successfully amplified, while the remaining three isolates yielded multiple bands. Among the *M. nivale* isolates, only two out of seven were successfully amplified; the remaining isolates failed to amplify under the conditions tested. Two *M. majus* isolates were sent for sequencing in the forward direction (LSD) and yielded partial gene sequences that were 881 (isolate 99027) and 888 bp (isolate 99049) in length. For both sequences, the top match when searched against the GenBank database (blastn) was the β-tubulin gene from *Pestalotiopsis paeniicola* (Order Xylariales, Family Amphisphaeriaceae; accession FJ975603.1). These *M. majus* sequences were then aligned with 19 β-tubulin sequences (Table

39
2.4) to design new primers. Following optimization, this primer set (Btub526F / Btub1332R) successfully amplified a sequence of the predicted length from all of the *M. nivale* and *M. majus* isolates with which they were tested.

**2.3.2 Sequence differences between *Microdochium nivale* and *M. majus***

Following PCR protocol optimization and primer redesign as necessary, the genomic regions studied were successfully amplified in all of the isolates included in these experiments, with the exception described by Glynn et al. (2005) that *M. majus* isolates were only amplified by the EFMajF/EFMicR primer pair, and *M. nivale* isolates were only amplified by the EFNivF/EfMicR primer pair. These were specific primers designed to amplify the different taxa separately (Glynn et al. 2005). The Y13N primer set designed to amplify *M. nivale* but not *M. majus* (Nicholson et al. 1996) also failed to amplify some *M. nivale* isolates (described below).

All of the sequences generated in these analyses have been deposited in the NCBI GenBank database (accession numbers JX280526-JX280607).

The Y13N and Y13M primer sets were tested with a set of European and North American strains of *M. nivale* and *M. majus* (Table 2.1). Among the *M. majus* isolates tested, nine out of nine of the European isolates, and six out of six of the North American isolates tested were successfully amplified. Among the *M. nivale* isolates, 12 of 13 of the European isolates tested and 6 of 22 of the North American isolates tested were amplified (Figure 2.2).

For RPB2, the primer set RPB150F / fRPB2_7cR produced a single amplicon of 1,040 bp in length for all 18 of the samples tested (Table 2.1). These amplicons were sequenced in the forward direction only using the primer RPB150_F, yielding raw sequences that varied between 970-1180 bp in length. The alignments were performed using a 729 bp region common among
all of the sequences obtained (Appendix 2.3). There were 62 parsimony-informative characters, 581 constant characters, and 86 parsimony-uninformative characters. The six *M. majus* isolates grouped together into a single clade with bootstrap values of 100% in all three analyses performed. The 12 *M. nivale* isolates also grouped together into a single clade with bootstrap values of 100% in the MP and NJ trees, and 97% in the ML tree (Figure 2.3).

The *HindIII* digest of the RPB2 isolates confirmed the presence of these consistent sequence differences in all of the *M. nivale* and *M. majus* isolates tested. While *M. nivale* isolates were cut two times to produce fragments that were 125, 444, and 471 bp in length, the *M. majus* isolates were cut only once, producing fragments that were 444 and 596 bp in length (Figure 2.4). An alignment showing the cut sites for this restriction enzyme is found in Appendix 2.3.

For the β-tubulin sequences, a single amplicon of 767 bp was obtained for all 18 isolates tested using the primer pair Btub526F and Btub1332R, and these amplicons were sequenced with the forward primer yielding raw sequences that varied between 723 to 814 bp in length. The alignments were performed using a 720 bp fragment common among all of the sequences obtained (Appendix 2.4). There were 55 parsimony-informative characters, 651 constant characters, and 14 parsimony-uninformative characters. The *M. majus* isolates tested formed a single clade with bootstrap values of 100, 91, and 99% in the NJ, ML, and MP trees. The *M. nivale* isolates tested formed a single clade with 100% bootstrap support in the NJ and MP trees, and 98% in the ML tree (Figure 2.3).

For EF-1α, a single amplicon of either 491 (*M. nivale*) or 487 (*M. majus*) bp in length was obtained for all 18 isolates tested, and these amplicons were sequenced in the forward direction using either EFMajF for *M. majus* or EFNivF for *M. nivale* isolates. The raw sequences obtained ranged between 450 to 491 bp in length, and a 423 bp region common to all isolates
was aligned for analysis (Appendix 2.5). There were 4 parsimony-informative characters, 387 constant characters, and 38 parsimony-uninformative characters. The *M. majus* isolates tested formed a single group with bootstrap support of 87, 60, and 74 % in the NJ, ML, and MP trees. The *M. nivale* isolates tested formed a single group with bootstrap support of 92, 89, and 93% in NJ, ML, and MP trees (Figure 2.3).

For ITS, a single amplicon 530 bp in length was obtained for the 20 isolates tested, and these amplicons were sequenced in the forward direction using the primer ITS1. The raw sequences obtained ranged between 503 to 530 bp in length, and a 503 bp region common to all isolates was aligned for analysis (Appendix 2.6). There were 21 parsimony-informative characters, and 482 constant characters. All variable characters were parsimony-informative. Only the NJ analysis grouped the *M. majus* isolates (50% bootstrap support) and the *M. nivale* isolates (99% bootstrap support) into distinct clades (Figure 2.3).

### 2.3.3 Geographic and host-specific differences

For RPB2, the European *M. majus* isolates formed distinct clades, with NJ, ML, and MP bootstrap values of 99, 100, and 100%, respectively. The North American isolates did not form a single cluster in any of the analyses. Among the *M. nivale* isolates, the European and North American isolates did not form distinct clades, but all isolates from turf formed a group with bootstrap values of 88, 90, and 96%, as did the wheat isolates with bootstrap values of 85% in the NJ and 73% in both the ML and MP trees.

For β-tubulin, among the *M. nivale* isolates tested, all European and North American samples isolated from wheat formed a clade with 81, 73, and 77% bootstrap support in the NJ, ML, and MP trees. The *M. nivale* isolates from turf formed a single group with 53, 54, and 59%
bootstrap support in the NJ, ML, and MP trees. No distinction could be made between isolates from North America and those from Europe in any of the analyses with the partial β-tubulin sequences. Neither the ITS nor the partial EF-1α sequences resolved groups based on geographic or host origin for either *M. majus* or *M. nivale* but only a limited number of isolates were tested here. A greater number of representatives from each geographic or host origin might reveal some other relationships.

### 2.4 Discussion

In this Chapter, the nucleotide sequences of four genomic regions were used to explore the differences between and among the sister species *Microdochium nivale* and *M. majus*. Differences between European and North American isolates were investigated by studying an equal number of isolates from both continents. Within *M. nivale*, which has both a larger host range and a higher range of reported genetic diversity (Maurin et al. 1995), isolates from both turfgrasses and wheat were examined to assess whether differences due to host plant origin could be identified.

All three of the protein-coding regions examined (RPB2, β-tubulin, and EF-1α) resolved all of the *M. majus* and *M. nivale* isolates studied into separate clades with strong bootstrap support. These specific groupings were not resolved with ITS. The ITS region also failed to resolve any of the other clades supported by the other genomic regions studied. Despite the apparent inability of ITS to distinguish between *M. nivale* and *M. majus*, digestion of the ITS amplified region with *Rsa*I has been shown to produce a single cut for *M. majus* isolates, while *M. nivale* is uncut (Maurin et al. 1995). Examination of the sequences obtained in this study
found that the presence or absence of this restriction site as reported by other researchers was conserved in all of the European and North American isolates sequenced (Appendix 2.6).

The inability of ITS sequences to resolve *M. nivale* and *M. majus* into separate clades using various tree-building methods supports the assertion that ITS may not be useful for all species-level fungal phylogenies. Concerted evolution of the multiple copies of this sequence within each individual’s genome may slow the rate of change of this sequence and maintain conservation between closely-related taxa, although the presumed lack of selective pressure allows this region to accumulate mutations between more distantly-related groups (Elder 1995). Furthermore, due to the high copy number of ITS within a genome, direct sequencing of a PCR product without cloning may yield polymorphic results, with clear variability within the chromatogram reflecting site ambiguity. The sequence thus obtained is essentially a consensus sequence that probably reflects the most common of the variable characters, but which may also obscure real differences between sister species that have been separated for only a short evolutionary time.

The EF-1α region, used as direct evidence to support the elevation of *M. nivale* and *M. majus* to sister species by Glynn et al. (2005), resolved these taxa into distinct clades in this study. In the paper by Glynn and colleagues, two separate groups of primers were designed: the EF1/F and EF1/R set, which produced an amplicon of approximately 840 bp from both *M. nivale* and *M. majus*, and the EFNivF, EFMajF, and EFMicR primers, which, together, amplified fragments of approximately 430 bp from either *M. nivale* or *M. majus* exclusively, and a separate set of primers, (2005). This latter set of primers was used to produce a 430 bp fragment of EF-1α in this study. The amplicon contained only four parsimony-informative positions and did not resolve further sub-divisions, such as geographical regions. However, Glynn et al. (2005) used a
832 bp amplicon (produced by the EF1/F and EF1/R primers) of which 56 of the 77 phylogenetically informative positions they identified were found outside of the region sequenced in this study. It is thus possible that the trends observed for RPB2 and β-tubulin may also be supported by larger portions or the full EF-1α sequence. The EF-1α primers used in this study did not amplify DNA from *M. bolleyi* despite several attempts to optimize the reaction. This result suggests that nucleotide-level differences exist within the EF-1α regions of *Microdochium* species. This hypothesis is further explored in Chapter 3.

Among the *M. nivale* isolates, both RPB2 and β-tubulin supported the resolution of distinct clades for all of the isolates originally collected on wheat, regardless of their geographic origin. These genes also supported a distinct clade for isolates originally collected from bentgrasses (*Agrostis* spp.) in five out of the six analyses performed. None of the analyses supported a distinct clade for either the North American or the European *M. nivale* isolates. All *M. majus* isolates from North America were collected in one geographical area near Atwood, Ontario, but the β-tubulin tree demonstrates that these isolates were not identical in sequence. Despite the fact that the North American *M. majus* samples were collected in close proximity, none of the sequences were identical for any of the four genetic regions examined. Similarly, multiple isolates of *M. majus* were chosen for analysis (rather than the same three strains for all analyses) to examine the diversity of the *M. majus* isolates in our collection. The presence of genetic variation within *M. majus*, in addition to reports of perithecia (Parry et al. 1995), suggests that sexual recombination occurs for this species in the field.

Together, these results suggest that host specialization has occurred within *M. nivale*, which exhibits a wider host range than *M. majus*. This finding is concordant with the observation that *M. nivale* isolates collected across Europe from a variety of plant hosts were genetically
heterogeneous (Maurin et al. 1995). Similarly, the RFLP profiles of isolates of *M. nivale* sensu stricto from a single turfgrass host were more similar to those obtained for isolates from a different geographic location than they were to those of isolates from a different turfgrass species (Mahuku et al. 1998). High genetic diversity was also observed within the populations studied (Mahuku et al. 1998), suggesting that the differences observed between populations cannot be explained only by ecological separation. Sexual reproduction, which has not been explicitly observed in the field for *M. nivale* (Smith 1992), but which has been observed in vitro (Lees et al. 1995) and for which there is indirect evidence in the way of linkage disequilibrium calculations (Mahuku et al. 1998), may explain these observations.

For *M. majus*, the RPB2 analyses alone supported a group containing all of the European isolates studied, but the North American isolates were not grouped together. However, some differences were observed between the North American and European *M. nivale* isolates using the Y13NF/Y13NR primer set. Both the Y13N and the and Y13M primer sets are reported to discriminate between *M. nivale* and *M. majus* by selectively amplifying only *M. nivale* (Y13N) or *M. majus* (Y13M). Both sets of Y13 primers are RAPD primers amplifying genomic regions of unknown function. The original publication describing these primers found that the Y13N pair failed to amplify the single North American *M. nivale* isolate that was tested (Nicholson et al. 1996). All of the *M. majus* isolates tested in this study, regardless of their geographic origin, were successfully amplified by the Y13M primer set. However, 16 of the 22 North American *M. nivale* isolates tested failed to amplify with the Y13N primer set, whereas 12 of the 13 European isolates tested under the same conditions were amplified successfully. These results demonstrate not only the high level of sequence variability within *M. nivale*, but also the necessity of testing
screening primers of this type with the widest variety of isolates possible to ensure that consistent results can be obtained. This problem is further explored in Chapter 3.

Combined with the results from other analyses, these data suggest that there is evidence for geographic specialization within both *M. nivale* and *M. majus*. For *M. nivale*, host species specialization may have a larger effect on genetic diversity than geographic location. To address this hypothesis, whole-genome sequencing of *M. majus* isolates from a wider range of plant hosts would be informative. Relative to *M. nivale*, *M. majus* has previously displayed a small amount of genetic heterogeneity regardless of host plant origin (Maurin et al. 1995), so the apparent lack of geographic specialization is in concordance with these observations.

Among the four genetic regions studied, β-tubulin and RPB2 were found to be more phylogenetically informative than either ITS or EF-1α, all four of which are commonly used in multi-gene phylogenies. β-tubulin has previously been used to study closely related sister species (Myllys 2001), and resolved *M. nivale* and *M. majus* into distinct clades in this study. In a recent multi-gene analysis that discovered cryptic species within the *Neofusicoccum parvum/N. ribis* species complex, RPB2 was also found to contain the largest number of phylogenetically informative characters (Pavlic 2009), which is also consistent with the 2009 assertion of Scoch et al. that, among a set of six genes (three protein-coding and three rDNA sequences), RPB2 was the most phylogenetically useful gene studied.

Overall, the data described in this Chapter support the reclassification of *M. nivale* and *M. majus* as distinct species. The hypothesis that there may be distinct sub-groups within *M. nivale* and *M. majus* were also supported, but additional isolates of each species should be examined before the presence or absence of distinct varieties within these species can be determined. To further explore the level of variation within and between *M. nivale* and *M. majus*,
the whole-genome sequences of three isolates (two *M. nivale* and one *M. majus*) were obtained (Chapter 3).
2.5 References for Chapter 2


Congress Melbourne, Australia, July 2011. International Association for Plant Taxonomy, Bratislava.


Table 2.1 Isolates of *Microdochium nivale* and *M. majus*, including their geographic origin and host-plant origin used for nucleotide sequence analysis.

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Species</th>
<th>Geographic origin</th>
<th>Host plant</th>
<th>Genomic Regions Studied*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RPB2</td>
</tr>
<tr>
<td>99027</td>
<td><em>M. majus</em></td>
<td>Atwood, Canada</td>
<td><em>Triticum</em> sp.</td>
<td>-</td>
</tr>
<tr>
<td>99049</td>
<td><em>M. majus</em></td>
<td>Atwood, Canada</td>
<td><em>Triticum</em> sp.</td>
<td>+</td>
</tr>
<tr>
<td>99061</td>
<td><em>M. majus</em></td>
<td>Atwood, Canada</td>
<td><em>Triticum</em> sp.</td>
<td>+</td>
</tr>
<tr>
<td>99064</td>
<td><em>M. majus</em></td>
<td>Atwood, Canada</td>
<td><em>Triticum</em> sp.</td>
<td>+</td>
</tr>
<tr>
<td>12043</td>
<td><em>M. majus</em></td>
<td>Ottawa, Canada</td>
<td><em>Triticum</em> sp.</td>
<td>-</td>
</tr>
<tr>
<td>12044</td>
<td><em>M. majus</em></td>
<td>Ottawa, Canada</td>
<td><em>Triticum</em> sp.</td>
<td>-</td>
</tr>
<tr>
<td>12045</td>
<td><em>M. majus</em></td>
<td>Ottawa, Canada</td>
<td><em>Triticum</em> sp.</td>
<td>-</td>
</tr>
<tr>
<td>12046</td>
<td><em>M. majus</em></td>
<td>Ottawa, Canada</td>
<td><em>Triticum</em> sp.</td>
<td>-</td>
</tr>
<tr>
<td>10095</td>
<td><em>M. majus</em></td>
<td>Argilly, France</td>
<td><em>Triticum</em> sp.</td>
<td>-</td>
</tr>
<tr>
<td>10096</td>
<td><em>M. majus</em></td>
<td>Argilly, France</td>
<td><em>Triticum</em> sp.</td>
<td>-</td>
</tr>
<tr>
<td>10097</td>
<td><em>M. majus</em></td>
<td>Argilly, France</td>
<td><em>Triticum</em> sp.</td>
<td>-</td>
</tr>
<tr>
<td>10098</td>
<td><em>M. majus</em></td>
<td>Argilly, France</td>
<td><em>Triticum</em> sp.</td>
<td>+</td>
</tr>
<tr>
<td>10099</td>
<td><em>M. majus</em></td>
<td>Medicina, Italy</td>
<td><em>Triticum</em> sp.</td>
<td>+</td>
</tr>
<tr>
<td>10100</td>
<td><em>M. majus</em></td>
<td>Medicina, Italy</td>
<td><em>Triticum</em> sp.</td>
<td>-</td>
</tr>
<tr>
<td>10148</td>
<td><em>M. majus</em></td>
<td>Castelnaudary, France</td>
<td><em>Triticum</em> sp.</td>
<td>-</td>
</tr>
<tr>
<td>10149</td>
<td><em>M. majus</em></td>
<td>Bullion, France</td>
<td><em>Triticum</em> sp.</td>
<td>+</td>
</tr>
<tr>
<td>10150</td>
<td><em>M. majus</em></td>
<td>Aryn, France</td>
<td><em>Triticum</em> sp.</td>
<td>-</td>
</tr>
<tr>
<td>96101</td>
<td><em>M. nivale</em></td>
<td>Cambridge, Canada</td>
<td><em>Agrostis palustris</em></td>
<td>+</td>
</tr>
<tr>
<td>96103</td>
<td><em>M. nivale</em></td>
<td>Cambridge, Canada</td>
<td><em>Agrostis palustris</em></td>
<td>+</td>
</tr>
<tr>
<td>96107</td>
<td><em>M. nivale</em></td>
<td>Cambridge, Canada</td>
<td><em>Agrostis palustris</em></td>
<td>+</td>
</tr>
<tr>
<td>10085</td>
<td><em>M. nivale</em></td>
<td>Guelph, Canada</td>
<td><em>Agrostis palustris</em></td>
<td>-</td>
</tr>
<tr>
<td>11036</td>
<td><em>M. nivale</em></td>
<td>Guelph, Canada</td>
<td><em>Poa pratensis</em></td>
<td>-</td>
</tr>
<tr>
<td>12049</td>
<td><em>M. nivale</em></td>
<td>Guelph, Canada</td>
<td><em>Poa pratensis</em></td>
<td>-</td>
</tr>
</tbody>
</table>

*Note: The '+' and '-' symbols indicate whether the genomic region was amplified or not.*
<table>
<thead>
<tr>
<th>Isolate</th>
<th>Species</th>
<th>Geographic origin</th>
<th>Host plant</th>
<th>Genomic Regions Studied</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RPB2</td>
</tr>
<tr>
<td>12150</td>
<td><em>M. nivale</em></td>
<td>Guelph, Canada</td>
<td><em>Poa pratensis</em></td>
<td>-</td>
</tr>
<tr>
<td>12051</td>
<td><em>M. nivale</em></td>
<td>Guelph, Canada</td>
<td><em>Poa pratensis</em></td>
<td>-</td>
</tr>
<tr>
<td>12134</td>
<td><em>M. nivale</em></td>
<td>Guelph, Canada</td>
<td><em>Lolium perenne</em></td>
<td>-</td>
</tr>
<tr>
<td>12135</td>
<td><em>M. nivale</em></td>
<td>Guelph, Canada</td>
<td><em>Lolium perenne</em></td>
<td>-</td>
</tr>
<tr>
<td>12136</td>
<td><em>M. nivale</em></td>
<td>Guelph, Canada</td>
<td><em>Lolium perenne</em></td>
<td>-</td>
</tr>
<tr>
<td>12137</td>
<td><em>M. nivale</em></td>
<td>Guelph, Canada</td>
<td><em>Lolium perenne</em></td>
<td>-</td>
</tr>
<tr>
<td>12138</td>
<td><em>M. nivale</em></td>
<td>Guelph, Canada</td>
<td><em>Lolium perenne</em></td>
<td>-</td>
</tr>
<tr>
<td>12139</td>
<td><em>M. nivale</em></td>
<td>Guelph, Canada</td>
<td><em>Lolium perenne</em></td>
<td>-</td>
</tr>
<tr>
<td>12140</td>
<td><em>M. nivale</em></td>
<td>Guelph, Canada</td>
<td><em>Lolium perenne</em></td>
<td>-</td>
</tr>
<tr>
<td>12142</td>
<td><em>M. nivale</em></td>
<td>Atwood, Canada</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>12143</td>
<td><em>M. nivale</em></td>
<td>Atwood, Canada</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>12144</td>
<td><em>M. nivale</em></td>
<td>Atwood, Canada</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>12145</td>
<td><em>M. nivale</em></td>
<td>Ottawa, Canada</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>12146</td>
<td><em>M. nivale</em></td>
<td>Ottawa, Canada</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>12151</td>
<td><em>M. nivale</em></td>
<td>Ottawa, Canada</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>12152</td>
<td><em>M. nivale</em></td>
<td>Ottawa, Canada</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>12153</td>
<td><em>M. nivale</em></td>
<td>Ottawa, Canada</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>12154</td>
<td><em>M. nivale</em></td>
<td>Ottawa, Canada</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>12155</td>
<td><em>M. nivale</em></td>
<td>Ottawa, Canada</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>10082</td>
<td><em>M. nivale</em></td>
<td>UK</td>
<td><em>Agrostis sp.</em></td>
<td>+</td>
</tr>
<tr>
<td>10083</td>
<td><em>M. nivale</em></td>
<td>UK</td>
<td><em>Agrostis sp.</em></td>
<td>+</td>
</tr>
<tr>
<td>10101</td>
<td><em>M. nivale</em></td>
<td>St. Leon-Rot, Germany</td>
<td><em>Agrostis sp.</em></td>
<td>+</td>
</tr>
<tr>
<td>10102</td>
<td><em>M. nivale</em></td>
<td>Switzerland</td>
<td><em>Agrostis sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>10103</td>
<td><em>M. nivale</em></td>
<td>Ottobeuren, Germany</td>
<td><em>Agrostis sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>10104</td>
<td><em>M. nivale</em></td>
<td>Netherlands</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>10105</td>
<td><em>M. nivale</em></td>
<td>Germany</td>
<td><em>Agrostis sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>Isolate</td>
<td>Species</td>
<td>Geographic origin</td>
<td>Host plant</td>
<td>Genomic Regions Studied*</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>------------------------</td>
<td>--------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RPB2</td>
</tr>
<tr>
<td>99006</td>
<td><em>M. nivale</em></td>
<td>Atwood, Canada</td>
<td><em>Triticum sp.</em></td>
<td>+</td>
</tr>
<tr>
<td>99010</td>
<td><em>M. nivale</em></td>
<td>Atwood, Canada</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>99063</td>
<td><em>M. nivale</em></td>
<td>Atwood, Canada</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>99069</td>
<td><em>M. nivale</em></td>
<td>Atwood, Canada</td>
<td><em>Triticum sp.</em></td>
<td>+</td>
</tr>
<tr>
<td>99077</td>
<td><em>M. nivale</em></td>
<td>Atwood, Canada</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>99084</td>
<td><em>M. nivale</em></td>
<td>Atwood, Canada</td>
<td><em>Triticum sp.</em></td>
<td>+</td>
</tr>
<tr>
<td>10106</td>
<td><em>M. nivale</em></td>
<td>Medicina, Italy</td>
<td><em>Triticum sp.</em></td>
<td>+</td>
</tr>
<tr>
<td>10107</td>
<td><em>M. nivale</em></td>
<td>Medicina, Italy</td>
<td><em>Triticum sp.</em></td>
<td>+</td>
</tr>
<tr>
<td>10151</td>
<td><em>M. majus</em></td>
<td>Aryon, France</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>10152</td>
<td><em>M. nivale</em></td>
<td>Castelnaudary, France</td>
<td><em>Triticum sp.</em></td>
<td>+</td>
</tr>
<tr>
<td>10153</td>
<td><em>M. nivale</em></td>
<td>Bullion, France</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>10154</td>
<td><em>M. nivale</em></td>
<td>Aryon, France</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>10155</td>
<td><em>M. nivale</em></td>
<td>Verdun sur Doubs, France</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
<tr>
<td>10156</td>
<td><em>M. nivale</em></td>
<td>Corgoloin, France</td>
<td><em>Triticum sp.</em></td>
<td>-</td>
</tr>
</tbody>
</table>

\* RPB2 = RNA polymerase II; Btub = β-tubulin; EF1α = elongation factor-1α; ITS = rDNA internal transcribed spacer; Y13 = region of unknown function identified by RAPD analyses (Nicholson et al. 1996). A "+" indicates that the sequence of a given genetic region was analysed for a particular isolate, and "-" indicates that this genetic region was not analysed for this isolate.
Table 2.2 Primers used in PCR and sequencing reactions.

<table>
<thead>
<tr>
<th>Genetic Region</th>
<th>Amplicon Size (bp)</th>
<th>Primer Name</th>
<th>Primer Sequence (5’-3’)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITS</td>
<td>530</td>
<td>ITS1</td>
<td>TCCCCCTCGCTTTATGATATG</td>
<td>(White et al. 1990)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ITS4</td>
<td>TCCTCCGCTTTATGATATG</td>
<td>(White et al. 1990)</td>
</tr>
<tr>
<td>EF-1α</td>
<td>487* or 491†</td>
<td>EFMajF</td>
<td>CCCCTTCTCCCTATCGC</td>
<td>(Glynn et al. 2005)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EFNivF</td>
<td>GTCCCCCTGTCTGACTGTTGT</td>
<td>(Glynn et al. 2005)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EFMicR</td>
<td>TCGATGGAGTCGATGG</td>
<td>(Glynn et al. 2005)</td>
</tr>
<tr>
<td>RPB2</td>
<td>1,500</td>
<td>fRPB2-5F</td>
<td>GAYGAYMGWGATCAYTTYGG</td>
<td>(Liu et al. 1999)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>fRPB2-7cR</td>
<td>CCCATRGCTTGYTTRCCCAT</td>
<td>(Liu et al. 1999)</td>
</tr>
<tr>
<td></td>
<td>1,040‡</td>
<td>RPB150_F</td>
<td>CTGGGGWGATCARAAGAAGG</td>
<td>This study</td>
</tr>
<tr>
<td>β-tubulin</td>
<td>ca. 880</td>
<td>Bt2a</td>
<td>GGTAACCAAATCGGTGCTGTTTC</td>
<td>(Glass 1995)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bt1b</td>
<td>GACGAGATCGTTCATGTTGAACTC</td>
<td>(Glass 1995)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Btub526_F</td>
<td>CGAGCGYATGAGGYGTYACTTT</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Btub1332_R</td>
<td>TCATGTTCTTGGGGTACAAG</td>
<td>This study</td>
</tr>
<tr>
<td>Unknown</td>
<td>220§</td>
<td>Y13MF</td>
<td>CTTAGGGCGGAAGATCGC</td>
<td>(Nicholson et al. 1996)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Y13MR</td>
<td>ATCCCTTTCCGAGGTTG</td>
<td>(Nicholson et al. 1996)</td>
</tr>
<tr>
<td>Unknown</td>
<td>300‖</td>
<td>Y13NF</td>
<td>CCAGCCGATTTGTTGATTG</td>
<td>(Nicholson et al. 1996)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Y13NR</td>
<td>GGTCAACGAGGCGAGGTTG</td>
<td>(Nicholson et al. 1996)</td>
</tr>
</tbody>
</table>

* for EFMajF and EFMicR when paired with *M. majus* isolates only
† for EFNivF and EFMicR when paired with *M. nivale* isolates only
‡ when paired with fRPB2-7cR
§ when paired with *M. majus* only
‖ when paired with *M. nivale* only
Table 2.3 List of RPB2 sequences used to design primers for *M. nivale* and *M. majus* (Sordariomycetes, Xylariales) with taxonomic information and GenBank accession numbers

<table>
<thead>
<tr>
<th>Class</th>
<th>Order</th>
<th>Family</th>
<th>Genus and Species</th>
<th>GenBank Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eurotiomycetes</td>
<td>Eurotiales</td>
<td>Trichocomaceae</td>
<td><em>Penicillium chrysogenum</em></td>
<td>XM_002568249.1</td>
</tr>
<tr>
<td>Lecanoromycetes</td>
<td>Lecanorales</td>
<td>Ophioparmaceae</td>
<td><em>Ophioparma lapponica</em></td>
<td>DQ973089.1</td>
</tr>
<tr>
<td>Leotiomycetes</td>
<td>Helotiales</td>
<td>Sclerotiniaceae</td>
<td><em>Botrytis ricini</em></td>
<td>GQ860997.1</td>
</tr>
<tr>
<td>Saccharomycetes</td>
<td>Saccharomycetales</td>
<td>Saccharomycetaceae</td>
<td><em>Candida tropicalis</em></td>
<td>AY485615.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>Debaryomyces hansenii</em></td>
<td>XM_002770548.1</td>
</tr>
<tr>
<td>Sordariomycetes</td>
<td>Hypocreales</td>
<td>Peethambara spirostriata</td>
<td>EU710770.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Clavicipitaceae</td>
<td><em>Metarhizium anisopliae</em></td>
<td>FJ787323.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hypocreaceae</td>
<td><em>Arachnocrea stipata</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>Hypocrea lixii</em></td>
<td>FJ179608.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>H. voglmayrii</em></td>
<td>FJ179622.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Protocrea farinosa</td>
<td>EU703942.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>P. pallida</em></td>
<td>EU703944.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sphaerostilbella aureonitens</td>
<td>FJ442763.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sporophagomyces chrysostomus</td>
<td>EU710780.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cladobotryum cubitense</td>
<td>EU710771.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fusarium virguliforme</td>
<td>GU170599.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stachybotrys echinata</td>
<td>EF692518.1</td>
<td></td>
</tr>
</tbody>
</table>

57
<table>
<thead>
<tr>
<th>Class</th>
<th>Order</th>
<th>Family</th>
<th>Genus and Species</th>
<th>GenBank Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>Trichoderma ovalisporum</em></td>
<td>FJ442796.1</td>
</tr>
<tr>
<td>Sordariomycetes</td>
<td>Xylariales</td>
<td>Xylariaceae</td>
<td><em>Daldinia concentrica</em></td>
<td>DQ368651.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>Seynesia erumpens</em></td>
<td>AY641073.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>Xylaria hypoxylon</em></td>
<td>DQ368652.1</td>
</tr>
</tbody>
</table>
**Table 2.4** List of species used to design β-tubulin primers for *M. nivale* and *M. majus* with GenBank accession numbers. Other than the *Microdochium* species, all species included are members of the Xylariaceae.

<table>
<thead>
<tr>
<th>Genus and Species</th>
<th>GenBank Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Podosordaria mexicana</em></td>
<td>GQ502719.1</td>
</tr>
<tr>
<td><em>Xylaria escharoidea</em></td>
<td>GQ502709.1</td>
</tr>
<tr>
<td><em>Amphirosellinia fushanensis</em></td>
<td>GQ495950.1</td>
</tr>
<tr>
<td><em>Astrocystis bambusae</em></td>
<td>GQ495942.1</td>
</tr>
<tr>
<td><em>Stilbohypoxylon elaeicola</em></td>
<td>GQ495933.1</td>
</tr>
<tr>
<td><em>Discoxylaria myrmecophila</em></td>
<td>GQ487710.1</td>
</tr>
<tr>
<td><em>Penzigia cantareirensis</em></td>
<td>GQ478220.1</td>
</tr>
<tr>
<td><em>Kretzschmaria guyanensis</em></td>
<td>GQ478214.1</td>
</tr>
<tr>
<td><em>Entoleuca mammta</em></td>
<td>GQ470230.1</td>
</tr>
<tr>
<td><em>Rosellinia merrillii</em></td>
<td>GQ470229.1</td>
</tr>
<tr>
<td><em>Nemania macrocarpa</em></td>
<td>GQ470226.1</td>
</tr>
<tr>
<td><em>Hypoxylon investiens</em></td>
<td>FJ185299.1</td>
</tr>
<tr>
<td><em>Daldinia concentrica</em></td>
<td>FJ185285.1</td>
</tr>
<tr>
<td><em>Annulohypoxylon cohaerens</em></td>
<td>FJ185283.1</td>
</tr>
<tr>
<td><em>Whalleya microplaca</em></td>
<td>EF025614.1</td>
</tr>
<tr>
<td><em>Theissenia cinerea</em></td>
<td>EF025613.1</td>
</tr>
<tr>
<td><em>Kretzschmaria lucidula</em></td>
<td>EF025610.1</td>
</tr>
<tr>
<td><em>Nemania illita</em></td>
<td>EF025608.1</td>
</tr>
<tr>
<td><em>Creosphaeria sassafras</em></td>
<td>DQ840094.1</td>
</tr>
<tr>
<td><em>Microdochium majus</em></td>
<td>-</td>
</tr>
<tr>
<td><em>Microdochium majus</em></td>
<td>-</td>
</tr>
</tbody>
</table>

* In-house *M. majus* isolate 99027
† In-house *M. majus* isolate 99049
Figure 2.1 Number of research journal papers published between 2000-2012 including the terms "maximum likelihood", "maximum parsimony", and either "neighbor joining" or "neighbour joining" obtained by searching the google scholar database using the term phylog* and either "maximum likelihood", "maximum parsimony" or "'neighbor joining' OR 'neighbor joining'" for articles published between 2000 and 2012.
Figure 2.2 Gel image for European (lanes A-F) and North American (lanes G-I, K-O) isolates of *Microdochium nivale* amplified with the Y13NF and Y13NR primers, which target a genetic region of unknown function. All six of the European isolates (lanes A-F) were strongly amplified, whereas weak bands are present in only two (lanes H, I) of the North American samples. A negative control, containing water instead of DNA, was included in the reaction (lane Q). Select sizes (bp) from a 100 bp increment ladder (lane J) are indicated to the right.
Figure 2.3 Bootstrapped maximum likelihood trees for RPB2, β-tubulin, EF-1α, and ITS. The tips of the tree are labelled with isolate number, species (M or N for *M. majus* or *M. nivale*, respectively), host origin (W or T for wheat or turfgrass), and geographic origin (NA or EU for North America or Europe). Bootstrap values out of 100 are located on the respective branches. Nodes with less than 50% support were collapsed. Separate clades resolving the sister species and host-specific or geographic groupings are indicated by curly braces.
Figure 2.4 Gel image for HindIII digest of RPB2 amplicon of *M. nivale* (lanes A-C) and *M. majus* (lanes D-F). Partial RPB2 sequences were amplified with the primers RPB_150F and fRPB2-7cR. The amplicons of *M. nivale* isolates were digested at two locations to produce fragments that were 125, 444, and 475 bp in length and the *M. majus* amplicons were digested at only one location, producing fragments that were 444 and 596 bp in length. Note that the 444 and 475 bp bands were not well-resolved. A non-digested RPB2 amplicon (1,040 bp in length) from *M. majus* is included for comparison (lane G). Select sizes (bp) from a 100 bp increment ladder (lane J) are indicated to the right.
### Appendix 2.1 Alignment of RPB2 sequences from Sordariomycete species. Primer-binding sites are indicated by shading.

<table>
<thead>
<tr>
<th>Species</th>
<th>Primer-Binding Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Debaryomyces Hansenii</td>
<td>ATTTTACAAAAAGAAATTACCTAATATTACTCAAGAAGAAGGGTTTGA 50</td>
</tr>
<tr>
<td>Candida tropicalis</td>
<td></td>
</tr>
<tr>
<td>Protocrea pallida</td>
<td></td>
</tr>
<tr>
<td>Protocrea farinosa</td>
<td></td>
</tr>
<tr>
<td>Cladobotryum cubitense</td>
<td></td>
</tr>
<tr>
<td>Hypocreavergilayrii</td>
<td></td>
</tr>
<tr>
<td>Trichoderma ovalisporum</td>
<td></td>
</tr>
<tr>
<td>Hypocreaveriixii</td>
<td></td>
</tr>
<tr>
<td>Arachnocrea stipata</td>
<td></td>
</tr>
<tr>
<td>Sphaerostilbella aureonitens</td>
<td></td>
</tr>
<tr>
<td>Sporophagomyces chrysostomus</td>
<td></td>
</tr>
<tr>
<td>Metarhizium anisopilae</td>
<td></td>
</tr>
<tr>
<td>Stachybotrys echinata</td>
<td></td>
</tr>
<tr>
<td>Peethambara spirostriata</td>
<td></td>
</tr>
<tr>
<td>Fusarium virguliforme</td>
<td></td>
</tr>
<tr>
<td>Penicillium chrysogenum</td>
<td>TTATGCAGAAGGAATGCTGCCTCATATTTGCGAAGAAGGGAGTTGGA 49</td>
</tr>
<tr>
<td>Ophioparma lapponica</td>
<td></td>
</tr>
<tr>
<td>Seynesia erumpens</td>
<td></td>
</tr>
<tr>
<td>Daedalia concentrica</td>
<td></td>
</tr>
<tr>
<td>Xylaria hypoxylon</td>
<td></td>
</tr>
<tr>
<td>Botrytis ricini</td>
<td></td>
</tr>
<tr>
<td>Debaryomyces Hansenii</td>
<td>GACCCGCAAGGCTTTCTTTTAGGTTACATGTTACAGATTATGTTAT 100</td>
</tr>
<tr>
<td>Candida tropicalis</td>
<td></td>
</tr>
<tr>
<td>Protocrea pallida</td>
<td></td>
</tr>
<tr>
<td>Protocrea farinosa</td>
<td></td>
</tr>
<tr>
<td>Cladobotryum cubitense</td>
<td></td>
</tr>
<tr>
<td>Hypocreavergilayrii</td>
<td></td>
</tr>
<tr>
<td>Trichoderma ovalisporum</td>
<td></td>
</tr>
<tr>
<td>Hypocreaveriixii</td>
<td></td>
</tr>
<tr>
<td>Arachnocrea stipata</td>
<td></td>
</tr>
<tr>
<td>Sphaerostilbella aureonitens</td>
<td></td>
</tr>
<tr>
<td>Sporophagomyces chrysostomus</td>
<td></td>
</tr>
<tr>
<td>Metarhizium anisopilae</td>
<td></td>
</tr>
<tr>
<td>Stachybotrys echinata</td>
<td></td>
</tr>
<tr>
<td>Peethambara spirostriata</td>
<td></td>
</tr>
<tr>
<td>Fusarium virguliforme</td>
<td></td>
</tr>
<tr>
<td>Penicillium chrysogenum</td>
<td>AACCCGCAAGGCTTTCTTTTAGGTTACATGTTACAGATTATGTTAT 99</td>
</tr>
<tr>
<td>Ophioparma lapponica</td>
<td></td>
</tr>
<tr>
<td>Seynesia erumpens</td>
<td></td>
</tr>
<tr>
<td>Daedalia concentrica</td>
<td></td>
</tr>
<tr>
<td>Xylaria hypoxylon</td>
<td></td>
</tr>
<tr>
<td>Botrytis ricini</td>
<td></td>
</tr>
</tbody>
</table>

**RPBF: GAYGAYM6WGATCA1TTYSS**

<table>
<thead>
<tr>
<th>Species</th>
<th>Primer-Binding Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Debaryomyces Hansenii</td>
<td>GTGCCCTAGAAAGAAAAGAAACAGATGATGATGATCATTTTTGGTAAAAG 150</td>
</tr>
<tr>
<td>Candida tropicalis</td>
<td></td>
</tr>
<tr>
<td>Protocrea pallida</td>
<td></td>
</tr>
<tr>
<td>Protocrea farinosa</td>
<td></td>
</tr>
<tr>
<td>Cladobotryum cubitense</td>
<td></td>
</tr>
<tr>
<td>Hypocreavergilayrii</td>
<td></td>
</tr>
<tr>
<td>Trichoderma ovalisporum</td>
<td></td>
</tr>
<tr>
<td>Hypocreaveriixii</td>
<td></td>
</tr>
<tr>
<td>Arachnocrea stipata</td>
<td></td>
</tr>
<tr>
<td>Sphaerostilbella aureonitens</td>
<td></td>
</tr>
<tr>
<td>Sporophagomyces chrysostomus</td>
<td></td>
</tr>
<tr>
<td>Metarhizium anisopilae</td>
<td></td>
</tr>
<tr>
<td>Stachybotrys echinata</td>
<td></td>
</tr>
<tr>
<td>Peethambara spirostriata</td>
<td></td>
</tr>
<tr>
<td>Fusarium virguliforme</td>
<td></td>
</tr>
</tbody>
</table>

65
Penicillium chrysogenum
Stachybotrys echinata
Metarhizium anisopliae
Sporophagomyces chrysostomus
Sphaerostilbella aureonitens
Hypocrea lixi
Trichoderma ovalisporum
Cladobotryum cubitense
Protocrea pallida
Candida tropicalis
Debaryomyces hansenii
Seynesia erumpens
Botrytis cinerea

Debaryomyces hansenii
Candida tropicalis
Protocrea pallida
Protocrea farinosa
Cladobotryum cubitense
Hypocrea voglmayrii
Trichoderma ovalisporum
Hypocrea lixi
Arachnocrea stipata
Sphaerostilbella Aureonitens
Sporophagomyces chrysostomus
Metarhizium anisopliae
Stachybotrys echinata
Penicillium virguliforme
Penicillium chrysogenum
Ophioparma lapponica
Seynesia erumpens
Xylaria Hypoxylon
Botrytis cinerea

Debaryomyces hansenii
Candida tropicalis
Protocrea pallida
Protocrea farinosa
Cladobotryum cubitense
Hypocrea voglmayrii
Trichoderma ovalisporum
Hypocrea lixi
Arachnocrea stipata
Sphaerostilbella Aureonitens
Sporophagomyces chrysostomus
Metarhizium anisopliae
Stachybotrys echinata
Penicillium virguliforme
Penicillium chrysogenum
Ophioparma lapponica
Seynesia erumpens
Xylaria Hypoxylon
Botrytis cinerea

Debaryomyces hansenii
Candida tropicalis
Protocrea pallida
Protocrea farinosa
Cladobotryum cubitense
Hypocrea voglmayrii
Trichoderma ovalisporum
Hypocrea lixi
Arachnocrea stipata
Sphaerostilbella Aureonitens
Sporophagomyces chrysostomus
Metarhizium anisopliae
Stachybotrys echinata
Penicillium virguliforme
Penicillium chrysogenum
Ophioparma lapponica
Seynesia erumpens
Xylaria Hypoxylon
Botrytis cinerea

Debaryomyces hansenii
Candida tropicalis
Protocrea pallida
Protocrea farinosa
Cladobotryum cubitense
Hypocrea voglmayrii
Trichoderma ovalisporum
Hypocrea lixi
Arachnocrea stipata
Sphaerostilbella Aureonitens
Sporophagomyces chrysostomus
Metarhizium anisopliae
Stachybotrys echinata
Penicillium virguliforme
Penicillium chrysogenum
Ophioparma lapponica
Seynesia erumpens
Xylaria Hypoxylon
Botrytis cinerea

Debaryomyces hansenii
Candida tropicalis
Protocrea pallida
Protocrea farinosa
Cladobotryum cubitense
Hypocrea voglmayrii
Trichoderma ovalisporum
Hypocrea lixi
Arachnocrea stipata
Sphaerostilbella Aureonitens
Sporophagomyces chrysostomus
Metarhizium anisopliae
Stachybotrys echinata
Penicillium virguliforme
Penicillium chrysogenum
Ophioparma lapponica
Seynesia erumpens
Xylaria Hypoxylon
Botrytis cinerea

Debaryomyces hansenii
Candida tropicalis
Protocrea pallida
Protocrea farinosa
Cladobotryum cubitense
Hypocrea voglmayrii
Trichoderma ovalisporum
Hypocrea lixi
Arachnocrea stipata
Sphaerostilbella Aureonitens
Sporophagomyces chrysostomus
Metarhizium anisopliae
Stachybotrys echinata
Penicillium virguliforme
Penicillium chrysogenum
Ophioparma lapponica
Seynesia erumpens
Xylaria Hypoxylon
Botrytis cinerea
<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Penicillium chrysogenum</em></td>
<td>G---------CGGACAAAGA-CCTTGCACACGACACTGCGCAGAGAAC--------------------</td>
</tr>
<tr>
<td><em>Ophioparma lapponica</em></td>
<td>G---------ACGATTAGAGGCTCTGG-ACCAAAATATGGACGCGGAGAAGA----------------------------</td>
</tr>
<tr>
<td><em>Seynesia erumpens</em></td>
<td>G---------TGCACT---------CA----------GACC-CTCGCCCTGAGAACAGG-----------------</td>
</tr>
<tr>
<td><em>Daldinia concentrica</em></td>
<td>G---------CGAG---------CTGCTGACGTGAGGAGAAGAGC------------------------------</td>
</tr>
<tr>
<td><em>Xylaria hypoxylon</em></td>
<td>T---------ACGAT---------CA----------ATGCCTCCTCCGCCAGGAGT-------------------</td>
</tr>
<tr>
<td><em>Botrytis ricini</em></td>
<td>AAACACAGCTTTGAC---------CTTCGGTGTCTCTCCCAGGCAA-------------------------------</td>
</tr>
<tr>
<td></td>
<td>**</td>
</tr>
</tbody>
</table>
Penicillium chrysogenum
Ophioparma lapponica
Seynesia erumpens
Daldinia concentrica
Xylaria hypoxylon
Botrytis cinerea

Debaryomyces hansenii
Candida tropicalis
Protocrea pallida
Protocrea farinosa
Cladobotryum cubitense
Hypocrea voglmayrii
Trichoderma ovalisporum
Hypocrea lixii
Arachnocrea stipata
Sphaerostilbella aureonitens
Sporophagomyces chrysostomus
Metarhizium anisopliae
Stachybotrys echnata
Peethambara spirostriata
Fusarium virguliforme
Penicillium chrysogenum
Ophioparma lapponica
Seynesia erumpens
Daldinia concentrica
Xylaria hypoxylon
Botrytis cinerea

Debaryomyces hansenii
Candida tropicalis
Protocrea pallida
Protocrea farinosa
Cladobotryum cubitense
Hypocrea voglmayrii
Trichoderma ovalisporum
Hypocrea lixii
Arachnocrea stipata
Sphaerostilbella aureonitens
Sporophagomyces chrysostomus
Metarhizium anisopliae
Stachybotrys echnata
Peethambara spirostriata
Fusarium virguliforme
Penicillium chrysogenum
Ophioparma lapponica
Seynesia erumpens
Daldinia concentrica
Xylaria hypoxylon
Botrytis cinerea

Debaryomyces hansenii
Candida tropicalis
Protocrea pallida
Protocrea farinosa
Cladobotryum cubitense
Hypocrea voglmayrii
Trichoderma ovalisporum
Hypocrea lixii
Arachnocrea stipata
Sphaerostilbella aureonitens
Sporophagomyces chrysostomus
Metarhizium anisopliae
Stachybotrys echnata
Peethambara spirostriata
Fusarium virguliforme
Penicillium chrysogenum
Ophioparma lapponica
Seynesia erumpens
Daldinia concentrica
Xylaria hypoxylon
Botrytis cinerea

Debaryomyces hansenii
Candida tropicalis
Protocrea pallida
Protocrea farinosa
Cladobotryum cubitense
Hypocrea voglmayrii
Trichoderma ovalisporum
Hypocrea lixii
Arachnocrea stipata
Sphaerostilbella aureonitens
Sporophagomyces chrysostomus
Metarhizium anisopliae
Stachybotrys echnata
Peethambara spirostriata
Fusarium virguliforme
Penicillium chrysogenum
Ophioparma lapponica
Seynesia erumpens
Daldinia concentrica
Xylaria hypoxylon
Botrytis cinerea

Debaryomyces hansenii
Candida tropicalis
Protocrea pallida
Protocrea farinosa
Cladobotryum cubitense
Hypocrea voglmayrii
Trichoderma ovalisporum
Hypocrea lixii
Arachnocrea stipata
Sphaerostilbella aureonitens
Sporophagomyces chrysostomus
Metarhizium anisopliae
Stachybotrys echnata
Peethambara spirostriata
Fusarium virguliforme
Penicillium chrysogenum
Ophioparma lapponica
Seynesia erumpens
Daldinia concentrica
Xylaria hypoxylon
Botrytis cinerea

Debaryomyces hansenii
Candida tropicalis
Protocrea pallida
Protocrea farinosa
Cladobotryum cubitense
Hypocrea voglmayrii
Trichoderma ovalisporum
Hypocrea lixii
Arachnocrea stipata
Sphaerostilbella aureonitens
Sporophagomyces chrysostomus
Metarhizium anisopliae
Stachybotrys echnata
Peethambara spirostriata
Fusarium virguliforme
Penicillium chrysogenum
Ophioparma lapponica
Seynesia erumpens
Daldinia concentrica
Xylaria hypoxylon
Botrytis cinerea

Debaryomyces hansenii
Candida tropicalis
Protocrea pallida
Protocrea farinosa
Cladobotryum cubitense
Hypocrea voglmayrii
Trichoderma ovalisporum
Hypocrea lixii
Arachnocrea stipata
Sphaerostilbella aureonitens
Sporophagomyces chrysostomus
Metarhizium anisopliae
Stachybotrys echnata
Peethambara spirostriata
Fusarium virguliforme
Penicillium chrysogenum
Ophioparma lapponica
Seynesia erumpens
Daldinia concentrica
Xylaria hypoxylon
Botrytis cinerea

Debaryomyces hansenii
Candida tropicalis
Protocrea pallida
Protocrea farinosa
Cladobotryum cubitense
Hypocrea voglmayrii
Trichoderma ovalisporum
Hypocrea lixii
Arachnocrea stipata
Sphaerostilbella aureonitens
Sporophagomyces chrysostomus
Metarhizium anisopliae
Stachybotrys echnata
Peethambara spirostriata
Fusarium virguliforme
Penicillium chrysogenum
Ophioparma lapponica
Seynesia erumpens
Daldinia concentrica
Xylaria hypoxylon
Botrytis cinerea

Debaryomyces hansenii
Candida tropicalis
Protocrea pallida
Protocrea farinosa
Cladobotryum cubitense
Hypocrea voglmayrii
Trichoderma ovalisporum
Hypocrea lixii
Arachnocrea stipata
Sphaerostilbella aureonitens
Sporophagomyces chrysostomus
Metarhizium anisopliae
Stachybotrys echnata
Peethambara spirostriata
Fusarium virguliforme
Penicillium chrysogenum
Ophioparma lapponica
Seynesia erumpens
Daldinia concentrica
Xylaria hypoxylon
Botrytis cinerea
Penicillium chrysogenum
Ophioparma laponica
Seynesia erumpens
Daldinia concentrica
Xylaria hypoxylon
Botrytis ricini

---

**RPH1000 R (rev. c):**

Debaryomyces Hansenii

Candida tropicalis
Proteocrea palida
Proteocrea farinosa
Cladobotryum cubitense
Hypocrea voglmayrii
Trichoderma ovalisporum

Hypocrea lixi i

Arachnocrea stipata

Sphaerostilbella aureonitens

Sporophagomyces chrysostomus
Metarhizium anisopliae
Stachybotrys echinata
Fusarium virguliforme
Penicillium chrysogenum
Ophioparma laponica
Seynesia erumpens
Daldinia concentrica
Xylaria hypoxylon
Botrytis ricini

---

**AAARGGTTGGTTTTCTCGCAAAAGTTAGTTAATCTGCTACATAT**

74
Penicillium_chrysogenum  ------------------------TCGCCG--------CGTAACACC----TATCAGTGGG 1283
Ophioparma_lapponica  TATGTTCTATCGGCTGTTACCAACATAGCCGAAAATAGTAAGT 1161
Seynesia_erumpens  TGTGCT-------------------------CCCCCTTTCTATGGTAGAAT 1074
Daldinia_concentrica  ---------------------------------------------------------------
Xylaria_hypoxylon  ---------------------------------------------------------------
Botrytis_ricini  ---------------------------------------------------------------

RPB7CR (rev. c): 5'-ATGGGTAA--GCAAGCTATGGG-3'

Debaryomyces_hansenii  ATGGGTAA--GCAAGCTATGGGCTTTCTTGAGAGACTTCTTAGGAGA 1329
Candida_tropicalis  ATGGGTAA--GCAAGCTATGGGCTTTCTTGAGAGACTTCTTAGGAGA 1234
Protocrea_pallida  ---------------------------------------------------------------
Protocrea_farinosa  ---------------------------------------------------------------
Cladobotryum_cubitense  ---------------------------------------------------------------
Hypocrea_voglmayrii  ---------------------------------------------------------------
Trichoderma_ovalisporum  ---------------------------------------------------------------
Hypocrea_liti  ---------------------------------------------------------------
Arachnocrea_stipata  ---------------------------------------------------------------
Sphaerostilbella_aureonitens  ---------------------------------------------------------------
Sporophagomyces_chrysostomus  ---------------------------------------------------------------
Metarhizium_anisopliae  ---------------------------------------------------------------
Stachybotrys_echinata  ---------------------------------------------------------------
Peethambara_spirostriata  ---------------------------------------------------------------
Fusarium_virguliforme  ---------------------------------------------------------------
Penicillium_chrysogenum  ATGGGTAA--GCAAGCTATGGGCTTTCTTGAGAGACTTCTTAGGAGA 1330
Ophioparma_lapponica  CTCCTGCCA---AACACCTACCAGCTGCTGAGGTTTCCACATGGG 1207
Seynesia_erumpens  GTTTATTATCATCACTCTCTCTCTCTCTCTCTGAGGCTGAGGTTTCCACATGGG 1112
Daldinia_concentrica  ---------------------------------------------------------------
Xylaria_hypoxylon  ---------------------------------------------------------------
Botrytis_ricini  ---------------------------------------------------------------
Appendix 2.2 Alignment of β-tubulin sequences from Sordariomycete species. Priming sites are indicated by shading.

<table>
<thead>
<tr>
<th>Species</th>
<th>Priming Site</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entoleuca mammmta</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTTTCTAT</td>
<td>37</td>
</tr>
<tr>
<td>Rosellinia merzillii</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>46</td>
</tr>
<tr>
<td>Nemania macrocarpa</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>24</td>
</tr>
<tr>
<td>Nemania illita</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>22</td>
</tr>
<tr>
<td>Amphiroellina fushmanensis</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>21</td>
</tr>
<tr>
<td>Astrocytis balsasias</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>20</td>
</tr>
<tr>
<td>Kretzschmaria guyanensis</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>22</td>
</tr>
<tr>
<td>Kretzschmaria lucidula</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>25</td>
</tr>
<tr>
<td>Penzigia cantareirensis</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>20</td>
</tr>
<tr>
<td>Stilbohypoxylon elaeicola</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>24</td>
</tr>
<tr>
<td>Discoxylaria myrmecophila</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>26</td>
</tr>
<tr>
<td>Podosordaria mexicana</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>21</td>
</tr>
<tr>
<td>Whalleya microplaca</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>21</td>
</tr>
<tr>
<td>Creosphaeria sassafras</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>20</td>
</tr>
<tr>
<td>Annulohypoxylon cohaerens</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>20</td>
</tr>
<tr>
<td>Daldinia concentrica</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>22</td>
</tr>
<tr>
<td>Hypoxylon investiens</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>22</td>
</tr>
<tr>
<td>Theissenia cineans</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>21</td>
</tr>
<tr>
<td>Mn majus027</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>20</td>
</tr>
<tr>
<td>Mn majus049</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>20</td>
</tr>
<tr>
<td>Xylaria escharoidea</td>
<td>AAATAAGGTTGGATGTTGTAATTTTTTTTTTTTCTAT</td>
<td>21</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Species</th>
<th>Priming Site</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entoleuca mammmta</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>65</td>
</tr>
<tr>
<td>Rosellinia merzillii</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Nemania macrocarpa</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Nemania illita</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Amphiroellina fushmanensis</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Astrocytis balsasias</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Kretzschmaria guyanensis</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Kretzschmaria lucidula</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Penzigia cantareirensis</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Stilbohypoxylon elaeicola</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Discoxylaria myrmecophila</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Podosordaria mexicana</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Whalleya microplaca</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Creosphaeria sassafras</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Annulohypoxylon cohaerens</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Daldinia concentrica</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Hypoxylon investiens</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Theissenia cineans</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Mn majus027</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Mn majus049</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Xylaria escharoidea</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Species</th>
<th>Priming Site</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entoleuca mammmta</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Rosellinia merzillii</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Nemania macrocarpa</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Nemania illita</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Amphiroellina fushmanensis</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Astrocytis balsasias</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Kretzschmaria guyanensis</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Kretzschmaria lucidula</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Penzigia cantareirensis</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Stilbohypoxylon elaeicola</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Discoxylaria myrmecophila</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Podosordaria mexicana</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Whalleya microplaca</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Creosphaeria sassafras</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Annulohypoxylon cohaerens</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Daldinia concentrica</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
<tr>
<td>Hypoxylon investiens</td>
<td>TATCTGTGGCACTACTATATATATATATATATATATATATCT</td>
<td>74</td>
</tr>
</tbody>
</table>
Theissenia_cinerea

Entoleuca_mammta

Rosellinia_merrillii

Nemania_macrocarpa

Nemania_illita

Amphirosellinia_fusahenensis

Astrocytis_bambusae

Kretzschmaria_guyanensis

Penigla_cantareirensis

Discoxylaria_myrmecophila

Podosordaria_mexicana

Whalleya_microplaca

Creosphaeria_sassafiras

Anulohypoxylon_coaheraens

Daidinia_concentrica

Hypoxylon_investiens

Theissenia_cinerea

Nemania_macrocarpa

Kretzschmaria_guyanensis

Penigla_cantareirensis

Discoxylaria_myrmecophila

Podosordaria_mexicana

Whalleya_microplaca

Creosphaeria_sassafiras

Anulohypoxylon_coaheraens

Daidinia_concentrica

Hypoxylon_investiens

Theissenia_cinerea

Xylaria_escharoidea

Entoleuca_mammta

GAGTTA-------------------------GCTACGCA-TACCTTTCCC---GTCTCT 313
Rosellinia_merrillii

GAGTTA-------------------------GCTACGCA-TACCTTTCCC---GTCTCT 328
Nemania_macrocarpa

GAGTTA-------------------------GCTACGCA-TACCTTTCCC---GTCTCT 328
Nemania_illita

GAGTTA-------------------------GCTACGCA-TACCTTTCCC---GTCTCT 328
Amphirosellinia_fusahenensis

GAGTTA-------------------------GCTACGCA-TACCTTTCCC---GTCTCT 328
Astrocytis_bambusae

GAGTTA-------------------------GCTACGCA-TACCTTTCCC---GTCTCT 328
Kretzschmaria_guyanensis

GAGTTA-------------------------GCTACGCA-TACCTTTCCC---GTCTCT 328
Penigla_cantareirensis

GAGTTA-------------------------GCTACGCA-TACCTTTCCC---GTCTCT 328
Discoxylaria_myrmecophila

GAGTTA-------------------------GCTACGCA-TACCTTTCCC---GTCTCT 328
Podosordaria_mexicana

GAGTTA-------------------------GCTACGCA-TACCTTTCCC---GTCTCT 328
Whalleya_microplaca

GAGTTA-------------------------GCTACGCA-TACCTTTCCC---GTCTCT 328
Creosphaeria_sassafiras

GAGTTA-------------------------GCTACGCA-TACCTTTCCC---GTCTCT 328
Anulohypoxylon_coaheraens

GAGTTA-------------------------GCTACGCA-TACCTTTCCC---GTCTCT 328
Daidinia_concentrica

GAGTTA-------------------------GCTACGCA-TACCTTTCCC---GTCTCT 328
Hypoxylon_investiens

GAGTTA-------------------------GCTACGCA-TACCTTTCCC---GTCTCT 328
Theissenia_cinerea

GAGTTA-------------------------GCTACGCA-TACCTTTCCC---GTCTCT 328
Xylaria_escharoidea

GAGTTA-------------------------GCTACGCA-TACCTTTCCC---GTCTCT 328
<table>
<thead>
<tr>
<th>Genus/Species</th>
<th>Oligonucleotide Sequence(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Theissenia cinea</em></td>
<td>AAAATGCAA------------------AGCTAACCGCTGTCCTAGTGCTT--CTCTAGT 247</td>
</tr>
<tr>
<td><em>Mn_majus027</em></td>
<td>---------------------------------------------GCTAAC--CTCTAGTGCTT--CTCTAGT 107</td>
</tr>
<tr>
<td><em>Mn_majus049</em></td>
<td>---------------------------------------------GCTAAC--CTCTAGTGCTT--CTCTAGT 82</td>
</tr>
<tr>
<td><em>Xylaria escharoidea</em></td>
<td>ATACAAA------------------AGCTGACCTTGTTTTTTTCCGAGTGTG 283 ** **</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genus/Species</th>
<th>Oligonucleotide Sequence(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Entoleuca mamorta</em></td>
<td>AGTTCCACCTCTCAACGGCG--CAATGCTGAAG--CTTC--CTCTAGT 356</td>
</tr>
<tr>
<td><em>Rosellinia merrillii</em></td>
<td>AGTTCCACCTCTCAACGGCG--CAATGCTGAAG--CTTC--CTCTAGT 371</td>
</tr>
<tr>
<td><em>Nemania macrocarpa</em></td>
<td>AGTCCACCTCTCAACGGCG--CAATGCTGAAG--CTTC--CTCTAGT 371</td>
</tr>
<tr>
<td><em>Nemania ililota</em></td>
<td>AGTCCACCTCTCAACGGCG--CAATGCTGAAG--CTTC--CTCTAGT 371</td>
</tr>
<tr>
<td><em>Amphirosettia fushanensis</em></td>
<td>AGTCTTCAGCTCAACGGCG--CAATGCTGAAG--CTAC--CTCTAGT 352</td>
</tr>
<tr>
<td><em>Astrocytis bamusuae</em></td>
<td>AGTCTTCAGCTCAACGGCG--CAATGCTGAAG--CTAC--CTCTAGT 341</td>
</tr>
<tr>
<td><em>Kretzschmaria guyanensis</em></td>
<td>AGTCTTCAGCTCAACGGCG--CAATGCTGAAG--CTAC--CTCTAGT 327</td>
</tr>
<tr>
<td><em>Kretzschmaria lucidula</em></td>
<td>AGTCTTCAGCTCAACGGCG--CAATGCTGAAG--CTAC--CTCTAGT 321</td>
</tr>
<tr>
<td><em>Penziella cantareirensis</em></td>
<td>AGTCTTCAGCTCAACGGCG--CAATGCTGAAG--CTAC--CTCTAGT 321</td>
</tr>
<tr>
<td><em>Stilbohypoxylon elaiecicol</em></td>
<td>AGTCTTCAGCTCAACGGCG--CAATGCTGAAG--CTAC--CTCTAGT 321</td>
</tr>
<tr>
<td><em>Discomyelia myrmecophila</em></td>
<td>AGTCTTCAGCTCAACGGCG--CAATGCTGAAG--CTAC--CTCTAGT 321</td>
</tr>
<tr>
<td><em>Podosorium mexicana</em></td>
<td>AGTCTTCAGCTCAACGGCG--CAATGCTGAAG--CTAC--CTCTAGT 321</td>
</tr>
<tr>
<td><em>Whalleya microplaca</em></td>
<td>AGTCTTCAGCTCAACGGCG--CAATGCTGAAG--CTAC--CTCTAGT 321</td>
</tr>
<tr>
<td><em>Creosphaeria sassafras</em></td>
<td>AGTCTTCAGCTCAACGGCG--CAATGCTGAAG--CTAC--CTCTAGT 321</td>
</tr>
<tr>
<td><em>Annuohypoxylon cohaerens</em></td>
<td>AGTCTTCAGCTCAACGGCG--CAATGCTGAAG--CTAC--CTCTAGT 321</td>
</tr>
<tr>
<td><em>Daldinia concentrica</em></td>
<td>AGTCTTCAGCTCAACGGCG--CAATGCTGAAG--CTAC--CTCTAGT 321</td>
</tr>
<tr>
<td><em>Hypoxylon investiens</em></td>
<td>AGTCTTCAGCTCAACGGCG--CAATGCTGAAG--CTAC--CTCTAGT 321</td>
</tr>
<tr>
<td><em>Theissenia cinea</em></td>
<td>AGTCTTCAGCTCAACGGCG--CAATGCTGAAG--CTAC--CTCTAGT 321</td>
</tr>
<tr>
<td><em>Mn_majus027</em></td>
<td>AGG------------------CAAAACACTTCTGAGTTGCAAG--G---------CTCTAGA 141</td>
</tr>
<tr>
<td><em>Mn_majus049</em></td>
<td>AGG------------------CAAAACACTTCTGAGTTGCAAG--G---------CTCTAGA 116</td>
</tr>
<tr>
<td><em>Xylaria escharoidea</em></td>
<td>AGTCTTCAGCTCAACGGCG--CAATGCTGAAG--CTAC--CTCTAGT 321 ** ** ** ** **</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genus/Species</th>
<th>Oligonucleotide Sequence(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Entoleuca mamorta</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 378</td>
</tr>
<tr>
<td><em>Rosellinia merrillii</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 393</td>
</tr>
<tr>
<td><em>Nemania macrocarpa</em></td>
<td>CGCGTC-------------------CTGGACGAGTGGAAG--AG--TT---------------- 355</td>
</tr>
<tr>
<td><em>Nemania ililota</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 333</td>
</tr>
<tr>
<td><em>Amphirosettia fushanensis</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 359</td>
</tr>
<tr>
<td><em>Astrocytis bamusuae</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 349</td>
</tr>
<tr>
<td><em>Kretzschmaria guyanensis</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 349</td>
</tr>
<tr>
<td><em>Kretzschmaria lucidula</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 349</td>
</tr>
<tr>
<td><em>Penziella cantareirensis</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 349</td>
</tr>
<tr>
<td><em>Stilbohypoxylon elaiecicol</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 349</td>
</tr>
<tr>
<td><em>Discomyelia myrmecophila</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 349</td>
</tr>
<tr>
<td><em>Podosorium mexicana</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 349</td>
</tr>
<tr>
<td><em>Whalleya microplaca</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 337</td>
</tr>
<tr>
<td><em>Creosphaeria sassafras</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 297</td>
</tr>
<tr>
<td><em>Annuohypoxylon cohaerens</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 314</td>
</tr>
<tr>
<td><em>Daldinia concentrica</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 269</td>
</tr>
<tr>
<td><em>Hypoxylon investiens</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 269</td>
</tr>
<tr>
<td><em>Theissenia cinea</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 269</td>
</tr>
<tr>
<td><em>Mn_majus027</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 269</td>
</tr>
<tr>
<td><em>Mn_majus049</em></td>
<td>C-----TT-------------------CCGG--ATGACGCGGAG--GA--CT---------------- 127</td>
</tr>
<tr>
<td><em>Xylaria escharoidea</em></td>
<td>C-------------------ATGCTCTTAG--ATCTTT--ATACCTGCTT 346</td>
</tr>
</tbody>
</table>

**Bt2a:**

<table>
<thead>
<tr>
<th>Genus/Species</th>
<th>Oligonucleotide Sequence(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Entoleuca mamorta</em></td>
<td>---TCTGGTGG--ACTCACA--ATGAAAATCAT--AGGTAACCCAAATCGGTTGCG 420</td>
</tr>
<tr>
<td><em>Rosellinia merrillii</em></td>
<td>---TCTGGTGG--ACTCACA--ATGAAAATCAT--AGGTAACCCAAATCGGTTGCG 436</td>
</tr>
<tr>
<td><em>Nemania macrocarpa</em></td>
<td>---TCTGGTGG--ACTCACA--ATGAAAATCAT--AGGTAACCCAAATCGGTTGCG 397</td>
</tr>
<tr>
<td><em>Nemania ililota</em></td>
<td>---TCTGGTGG--ACTCACA--ATGAAAATCAT--AGGTAACCCAAATCGGTTGCG 397</td>
</tr>
<tr>
<td><em>Amphirosettia fushanensis</em></td>
<td>---TCTGGTGG--ACTCACA--ATGAAAATCAT--AGGTAACCCAAATCGGTTGCG 375</td>
</tr>
<tr>
<td><em>Astrocytis bamusuae</em></td>
<td>---TCTGGTGG--ACTCACA--ATGAAAATCAT--AGGTAACCCAAATCGGTTGCG 401</td>
</tr>
<tr>
<td><em>Kretzschmaria guyanensis</em></td>
<td>---TCTGGTGG--ACTCACA--ATGAAAATCAT--AGGTAACCCAAATCGGTTGCG 392</td>
</tr>
<tr>
<td><em>Kretzschmaria lucidula</em></td>
<td>---TCTGGTGG--ACTCACA--ATGAAAATCAT--AGGTAACCCAAATCGGTTGCG 386</td>
</tr>
<tr>
<td><em>Penziella cantareirensis</em></td>
<td>---TCTGGTGG--ACTCACA--ATGAAAATCAT--AGGTAACCCAAATCGGTTGCG 386</td>
</tr>
<tr>
<td><em>Stilbohypoxylon elaiecicol</em></td>
<td>---TCTGGTGG--ACTCACA--ATGAAAATCAT--AGGTAACCCAAATCGGTTGCG 465</td>
</tr>
<tr>
<td><em>Discomyelia myrmecophila</em></td>
<td>---TCTGGTGG--ACTCACA--ATGAAAATCAT--AGGTAACCCAAATCGGTTGCG 388</td>
</tr>
<tr>
<td><em>Podosorium mexicana</em></td>
<td>---TCTGGTGG--ACTCACA--ATGAAAATCAT--AGGTAACCCAAATCGGTTGCG 379</td>
</tr>
<tr>
<td><em>Whalleya microplaca</em></td>
<td>---TCTGGTGG--ACTCACA--ATGAAAATCAT--AGGTAACCCAAATCGGTTGCG 355</td>
</tr>
<tr>
<td><em>Creosphaeria sassafras</em></td>
<td>---TCTGGTGG--ACTCACA--ATGAAAATCAT--AGGTAACCCAAATCGGTTGCG 339</td>
</tr>
<tr>
<td><em>Annuohypoxylon cohaerens</em></td>
<td>---TCTGGTGG--ACTCACA--ATGAAAATCAT--AGGTAACCCAAATCGGTTGCG 316</td>
</tr>
<tr>
<td><em>Daldinia concentrica</em></td>
<td>---TCTGGTGG--ACTCACA--ATGAAAATCAT--AGGTAACCCAAATCGGTTGCG 317</td>
</tr>
<tr>
<td><em>Hypoxylon investiens</em></td>
<td>---TCTGGTGG--ACTCACA--ATGAAAATCAT--AGGTAACCCAAATCGGTTGCG 345</td>
</tr>
<tr>
<td>Genus</td>
<td>Accession</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Theissenia cinerea</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Rosellinia merrillii</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Nemania macrocarpa</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Nemania illita</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Amphirosetia_fushanensis</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Astrocystis bambovae</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Kretzschmaria MenuItem</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Penzigia cantareirensis</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Stilbohypoxylon elaeicola</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Discoryxylaria myrmecophila</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Podosordaria mexicana</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Whalleya microplaca</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Creosphaeria sassafras</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Annulohypoxylon cohaerens</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Daldinia concentrica</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Hypoxylon investiens</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Theissenia cinerea</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Mn majus027</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
<tr>
<td>Mn majus049</td>
<td>CAACAAATTTCGCCGACGAGCAGCCCTCGGACGCCATTGGGC</td>
</tr>
</tbody>
</table>

**Theissenia_cinerea**

---

**Annulohypoxylon_cohaerens**

---

**Discoxylaria_myrmecophila**

---

**Amphirosellinia_fushanensi**

---

**Nemania_macrocarpa**

---

**Mn_majus027**

---

**Xylaria_escharoidea**

---

**Entoleuca_mammta**

---

**Rosellinia_merrillii**

---

**Nemania_illita**

---

**Amphirosettinia_fushanensis**

---

**Astrocytia_bambusae**

---

**Kretzschmaria_guyanensis**

---

**Kretzschmaria_lucidula**

---

**Penzigia_cantareirensis**

---

**Stilbophyoxylon_elaeicola**

---

**Discoxylaria_myrmecophila**

---

**Podosordaria_mexicana**

---

**Whalleya_microplaca**

---

**Creosphaeria_sassafra**

---

**Annulohypoxylon_coaheren**

---

**Daldinia_concentrica**

---

**Hypoxylon_investiens**

---

**Nemania**

---

**Theissenia_cinerea**

---

**Annulohypoxylon_cohaerens**

---

**Creosphaeria_sassafras**

---

**Xylaria_escharoidea**

---

**Entoleuca_mammta**

---

**Rosellinia_merrillii**

---

**Nemania_macrocarpa**

---

**Nemania_illita**

---

**Amphirosettinia_fushanensis**

---

**Astrocytia_bambusae**

---

**Kretzschmaria_guyanensis**

---

**Kretzschmaria_lucidula**

---

**Penzigia_cantareirensis**

---

**Stilbophyoxylon_elaeicola**

---

**Discoxylaria_myrmecophila**

---

**Podosordaria_mexicana**

---

**Whalleya_microplaca**

---

**Creosphaeria_sassafra**

---

**Annulohypoxylon_coaheren**

---

**Daldinia_concentrica**

---

**Hypoxylon_investiens**

---

**Theissenia_cinerea**

---

**Annulohypoxylon_cohaerens**

---

**Creosphaeria_sassafras**

---

**Xylaria_escharoidea**

---

**Entoleuca_mammta**

---

**Rosellinia_merrillii**

---

**Nemania_macrocarpa**

---

**Nemania_illita**

---

**Amphirosettinia_fushanensis**

---

**Astrocytia_bambusae**

---

**Kretzschmaria_guyanensis**

---

**Kretzschmaria_lucidula**

---

**Penzigia_cantareirensis**

---

**Stilbophyoxylon_elaeicola**

---

**Discoxylaria_myrmecophila**

---

**Podosordaria_mexicana**

---

**Whalleya_microplaca**

---

**Creosphaeria_sassafra**

---

**Annulohypoxylon_coaheren**

---

**Daldinia_concentrica**

---

**Hypoxylon_investiens**

---

**Theissenia_cinerea**

---

**Annulohypoxylon_cohaerens**

---

**Creosphaeria_sassafras**

---

**Xylaria_escharoidea**

---

**Entoleuca_mammta**

---

**Rosellinia_merrillii**

---

**Nemania_macrocarpa**

---

**Nemania_illita**

---

**Annulohypoxylon_fushanensis**

---

**Astrocytia_bambusae**

---

**Kretzschmaria_guyanensis**

---

**Kretzschmaria_lucidula**

---

**Penzigia_cantareirensis**

---

**Stilbophyoxylon_elaeicola**

---

**Discoxylaria_myrmecophila**

---

**Podosordaria_mexicana**

---

**Whalleya_microplaca**

---

**Creosphaeria_sassafra**

---

**Annulohypoxylon_coaheren**

---

**Daldinia_concentrica**

---

**Hypoxylon_investiens**

---

**Theissenia_cinerea**

---

**Annulohypoxylon_cohaerens**

---

**Creosphaeria_sassafras**

---

**Xylaria_escharoidea**

---

**Entoleuca_mammta**

---

**Rosellinia_merrillii**

---

**Nemania_macrocarpa**

---

**Nemania_illita**

---

**Annulohypoxylon_fushanensis**

---

**Astrocytia_bambusae**

---

**Kretzschmaria_guyanensis**

---

**Kretzschmaria_lucidula**

---

**Penzigia_cantareirensis**

---

**Stilbophyoxylon_elaeicola**

---

**Discoxylaria_myrmecophila**

---

**Podosordaria_mexicana**

---

**Whalleya_microplaca**

---

**Creosphaeria_sassafra**

---

**Annulohypoxylon_coaheren**

---

**Daldinia_concentrica**

---

**Hypoxylon_investiens**

---

**Theissenia_cinerea**

---

**Annulohypoxylon_cohaerens**

---

**Creosphaeria_sassafras**

---

**Xylaria_escharoidea**

---

**Entoleuca_mammta**

---

**Rosellinia_merrillii**

---

**Nemania_macrocarpa**

---

**Nemania_illita**

---

**Annulohypoxylon_fushanensis**

---

**Astrocytia_bambusae**

---

**Kretzschmaria_guyanensis**

---

**Kretzschmaria_lucidula**

---

**Penzigia_cantareirensis**

---

**Stilbophyoxylon_elaeicola**

---

**Discoxylaria_myrmecophila**

---

**Podosordaria_mexicana**

---

**Whalleya_microplaca**

---

**Creosphaeria_sassafra**

---

**Annulohypoxylon_coaheren**

---

**Daldinia_concentrica**

---

**Hypoxylon_investiens**

---

**Theissenia_cinerea**

---

**Annulohypoxylon_cohaerens**

---

**Creosphaeria_sassafras**

---

**Xylaria_escharoidea**

---
Hyphoxylon_investiens
Theissenia_cinerea
Mn_majus049
Mn_majus027
Enpleuca_mammta
Rosellinia_merrillii
Neama_macrocarpa
Neama_illa
Amphirosellinia_fusenhanis
Astrocytis_bambusae
Kretzschmaria_guyanensis
Astrocystis_bambusae
Amphirosellinia_fishenhanis
Ruosellinia_macrocarpa
Rosellinia_merrillii
Entoleuca_mammta
Xylaria_escaroidea

** ** ** * *** ** ** ** ** ** ** ** ** ** 824
** ** ** *********** * *** ** ***** ** **

83
**Hypoxylon Investiens**
GCTGAGCTGTGGACACACGCCTCTTCTGTCCGTCGTCGAGGCTGGGGG...797

**Theissenia Cineza**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...826

**Mn Majus027**
GCCAGCTGTGGACACCCGGTCTGAGGCTCCGCTGAGGCTGGGGG...522

**Mn Majus049**
GCCAGCTGGTGGACACCCGGTCTGAGGCTCCGCTGAGGCTGGGGG...498

**Xylaria Escharoidea**
GCTGATATTGGTCGACAGGCCAGCTCCGCTGAGGCTGGGGG...845

* *** * *** * *** * *** * *** * *** * *** * *** * *** *

**Entoleuca Mammta**
CTCGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...924

**Rosellinia Merrillii**
CTCGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...946

**Nemania Macrocarpa**
CTCGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...911

**Nemania Illita**
CTCGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...878

**AmphiroSELLiina Fushanensis**
CTCGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...913

**Astrocytis Bambusae**
CTCGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...907

**Kretzschmaria Guyanensis**
CTCGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...903

**Kretzschmaria Lucidula**
CTCGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...890

**Penzigia Cantareirensis**
CTCGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...892

**Stilbophyoxylon Elaeicola**
CTCGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...890

**Discoxyliarla Myrmecophila**
CTCGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...882

**Podosordaria Mexicanana**
CTCGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...947

**Whalleya Microplaca**
CTCGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...931

**Annuhpyoxylon Coheraena**
ATGTGATTGGTCGACCAGGTCCTCGAGGTCGTCCCGCGAGGCTGAGGG...522

**Daldinia Concentrica**
GCTGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...817

**Hypoxylon Investiens**
TTCGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...847

**Theissenia Cineza**
CTCGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...876

**Mn Majus027**
CTCGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...572

**Mn Majus049**
CTCGAGCTGCTCTCCAGGATCCATTCCATCTCCACCTCCGCTGAGGCTGGGGG...548

**Xylaria Escharoidea**
GCTGATATTGGTCGACAGGCCAGCTCCGCTGAGGCTGGGGG...895

* *** * *** * *** * *** * *** * *** * *** * *** * *** *

**Entoleuca Mammta**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...968

**Rosellinia Merrillii**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...990

**Nemania Macrocarpa**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...955

**Nemania Illita**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...922

**AmphiroSELLiina Fushanensis**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...951

**Astrocytis Bambusae**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...947

**Kretzschmaria Guyanensis**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...934

**Kretzschmaria Lucidula**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...936

**Penzigia Cantareirensis**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...1014

**Stilbophyoxylon Elaeicola**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...926

**Whalleya Microplaca**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...991

**Creosphearia Sassafras**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...977

**Annuhpyoxylon Coheraena**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...862

**Daldinia Concentrica**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...861

**Hypoxylon Investiens**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...891

**Theissenia Cineza**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...920

**Mn Majus027**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...615

**Mn Majus049**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...591

**Xylaria Escharoidea**
CTGTCATGGTCGTCGAGGCTCCGCTTCTCCGCTGAGGCTGGGGG...945

* *** * *** * *** * *** * *** * *** * *** * *** * *** *

---

84
<table>
<thead>
<tr>
<th>Genus</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daldinia_concentrica</td>
<td>CCAGCTGGTCGAGAACTCCTGGAATCAGGCTCAGTGCAGTTATGCTGATCACAGGACAGGCTCAG</td>
</tr>
<tr>
<td>Hypoxylon_investiens</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Theissenia_cinerentia</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Mn_majus027</td>
<td>GACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Mn_majus049</td>
<td>GACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Xylaria_escharoidea</td>
<td>GACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Entoleuca_mammta</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Rosellinia_merrillii</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Nemania_macrocarpa</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Nemania_illita</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Ampyrostellina_fusahenensis</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Astrocytis_bambusae</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Kretzschmaria_guyanensis</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Kretzschmaria_lucidula</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Penzigia_cantareirensis</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Stilbophyoxylon_elaecola</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Discoryzaria_myrmecophila</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Podosordaria_mexicana</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Whalleya_microplaca</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Creospheeria_sassafalas</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Annuolophyoxylon_cohereana</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Daldinia_concentrica</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Hypoxylon_investiens</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Theissenia_cinerentia</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Mn_majus027</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Mn_majus049</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
<tr>
<td>Xylaria_escharoidea</td>
<td>TGACTGATGACAGATCTGATGATCACTGACATCAGTACCACCTTCTTCTACTGATGACAGGCTCAG</td>
</tr>
</tbody>
</table>

86
Annulohypoxylon_cohaerens
Creosphaeria_sassafras
Discoxylaria_myrmecophila
Stilbohypoxylon_elaeicola
Penzigia_cantareirensis
Kretzschmaria_lucidula
Astrocystis_bambusae
Amphirosellinia_fushanensi
Nemania_illita
Nemania_macrocarpa
Rosellinia_merrillii
Xylaria_escharoidea

Entoleuca_mammta
Rosellinia_merrillii
Nemania_macrocarpa
Nemania_illita
Amphirosellinia_fushanensis
Astrocytis_bambusae
Kretzschmaria_guyanensis

Daldinia_concentrica
Annulohypoxylon_cohaerens
Creosphaeria_sassafras
Whalleya_microplaca
Creosphaeria_sassafras
Annulohypoxylon_cohaerens

** *****

Entoleuca_mammta
Rosellinia_merrillii
Nemania_macrocarpa
Nemania_illita
Amphirosellinia_fushanensis
Penzigia_cantareirensis
Stilbohypoxylon_elaeicola

Daldinia_concentrica
Annulohypoxylon_cohaerens

** *****

Entoleuca_mammta
Rosellinia_merrillii
Nemania_macrocarpa
Nemania_illita
Amphirosellinia_fushanensis

Penzigia_cantareirensis
Stilbohypoxylon_elaeicola

Daldinia_concentrica
Annulohypoxylon_cohaerens

Xylaria_escharoidea

Entoleuca_mammta
Rosellinia_merrillii
Nemania_macrocarpa
Nemania_illita
Amphirosellinia_fushanensis

Penzigia_cantareirensis
Stilbohypoxylon_elaeicola

Daldinia_concentrica
Annulohypoxylon_cohaerens

Xylaria_escharoidea

87
Annulohypoxylon_cohaerens
Creosphaeria_sassafras
Whalleya_microplaca
Penzigia_cantareirensis
Kretzschmaria_lucidula
Astrocystis_bambusae
Amphirosellinia_fushanensis
Nemania_macrocarpa
Rosellinia_merrillii
Entoleuca_mammta
Xylaria_escharoidea

Entoleuca_mammta
---TTACTAAT---TTTACGAC---CATTAGC---TGTTGCAAG---1516
Rosellinia_merrillii
---TCTGCTAAT---TCGCCAGTC---TCAATGGC---TGTTGCAAG---1538
Nemania_macrocarpa
---TTACTAAT---TTTGGGAC---TGACGGCTTGCGAAG---1503
Nemania_illita
---TTCTGTAAT---CTTTGCGAC---TAATCCGGGGCAG---1472
AmphiroSELLinia_fushanensis
---TCTGCTAAT---CTTTGACAT---CTGCGAGCTGGGCAAG---1510
Astrocytis_bambusae
---TTCTGCTAAT---CCAGAGAT---TCTGCTGGGCGGCAAG---1497
Kretzschmaria_guyanensis
---TTCTGCTAAT---CCAGAGAT---TCTGCTGGGCGGCAAG---1498
Kretzschmaria_lucidula
---TTCTGCTAAT---CATTAGC---TTTACGAC---TGTTGCAAG---1420
Penzigia_cantareirensis
---TTCTGCTAAT---TTTACGAC---TGTTGCAAG---1487
Stilbohypoxylon_elaeicola
---TTTACGAC---TGTTGCAAG---1561
Discoxyliarla_myrmecophila
---TTTACGAC---TGTTGCAAG---1489
Podosordaria_mexicana
AGCGGCTGTCATACGTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCTGGACGGCT
<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Daldinia_concentrica</em></td>
<td>TCGTCCCTACTTCG----------------------TTGAGTGGATTCCCAAA</td>
<td>1520</td>
</tr>
<tr>
<td><em>Hypoxylon_investiens</em></td>
<td>TCGTCCCTACTTCGTAATGGATTCCCAACAACATCCAGA</td>
<td>1570</td>
</tr>
<tr>
<td><em>Theissenia_cinerea</em></td>
<td>TCGTCCCTACTTCGTAATGGATTCCCAACAACATCCAGA</td>
<td>1570</td>
</tr>
<tr>
<td><em>Mn_majus027</em></td>
<td>-------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td><em>Mn_majus049</em></td>
<td>-------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td><em>Xylaria_escharoidea</em></td>
<td>TCCTCGTACTTCGTATGGATTCCCAACAACATCCAGA</td>
<td>1643</td>
</tr>
</tbody>
</table>
Appendix 2.3 RPB2 alignment of *M. nivale* and *M. majus* sequences with HindIII digest sites indicated by shading.

<table>
<thead>
<tr>
<th>Accession</th>
<th>Sequence</th>
<th>HindIII Digest Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>99084M_W_NA</td>
<td>AGATGGTAACACTTCGGCAAAACCTCGCCAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
<tr>
<td>99094M_W_NA</td>
<td>AGATGGTAACACTGCGCAAACTCGGCAAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
<tr>
<td>10149M_W_EU</td>
<td>AGATGGTAACACTGCGCAAACTCGGCAAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
<tr>
<td>99094M_W_NA</td>
<td>AGATGGTAACACTGCGCAAACTCGGCAAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
<tr>
<td>99084M_W_NA</td>
<td>AGATGGTAACACTGCGCAAACTCGGCAAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
<tr>
<td>99069N_W_NA</td>
<td>AGATGGTAACACTGCGCAAACTCGGCAAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
<tr>
<td>10101N_T_EU</td>
<td>AGATGGTAACACTGCGCAAACTCGGCAAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
<tr>
<td>10082N_T_EU</td>
<td>AGATGGTAACACTGCGCAAACTCGGCAAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
<tr>
<td>10083N_T_EU</td>
<td>AGATGGTAACACTGCGCAAACTCGGCAAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
<tr>
<td>96107N_T_NA</td>
<td>AGATGGTAACACTGCGCAAACTCGGCAAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
<tr>
<td>96101N_T_NA</td>
<td>AGATGGTAACACTGCGCAAACTCGGCAAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
<tr>
<td>96103N_T_NA</td>
<td>AGATGGTAACACTGCGCAAACTCGGCAAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
<tr>
<td>99061M_W_NA</td>
<td>AGATGGTAACACTGCGCAAACTCGGCAAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
<tr>
<td>10098M_W_EU</td>
<td>AGATGGTAACACTGCGCAAACTCGGCAAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
<tr>
<td>10099M_W_EU</td>
<td>AGATGGTAACACTGCGCAAACTCGGCAAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
<tr>
<td>10149M_W_EU</td>
<td>AGATGGTAACACTGCGCAAACTCGGCAAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
<tr>
<td>99049M_W_NA</td>
<td>AGATGGTAACACTGCGCAAACTCGGCAAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
<tr>
<td>99064M_W_NA</td>
<td>AGATGGTAACACTGCGCAAACTCGGCAAAATGGGCTTATGCTGTC 119</td>
<td></td>
</tr>
</tbody>
</table>

**HindIII digest site:**

```
<table>
<thead>
<tr>
<th>Accession</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>99064M_W_NA</td>
<td>TGGGACAGCTCCCGACAGCGGCTGCGCCGCTCGCTGCAAGACCTTCTCTGCTCGGTC 179</td>
</tr>
<tr>
<td>99049M_W_NA</td>
<td>TGGGACAGCTCCCGACAGCGGCTGCGCCGCTCGCTGCAAGACCTTCTCTGCTCGGTC 179</td>
</tr>
<tr>
<td>10149M_W_EU</td>
<td>TGGGACAGCTCCCGACAGCGGCTGCGCCGCTCGCTGCAAGACCTTCTCTGCTCGGTC 179</td>
</tr>
<tr>
<td>99064M_W_NA</td>
<td>TGGGACAGCTCCCGACAGCGGCTGCGCCGCTCGCTGCAAGACCTTCTCTGCTCGGTC 179</td>
</tr>
<tr>
<td>99049M_W_NA</td>
<td>TGGGACAGCTCCCGACAGCGGCTGCGCCGCTCGCTGCAAGACCTTCTCTGCTCGGTC 179</td>
</tr>
<tr>
<td>99064M_W_NA</td>
<td>TGGGACAGCTCCCGACAGCGGCTGCGCCGCTCGCTGCAAGACCTTCTCTGCTCGGTC 179</td>
</tr>
<tr>
<td>99049M_W_NA</td>
<td>TGGGACAGCTCCCGACAGCGGCTGCGCCGCTCGCTGCAAGACCTTCTCTGCTCGGTC 179</td>
</tr>
<tr>
<td>99064M_W_NA</td>
<td>TGGGACAGCTCCCGACAGCGGCTGCGCCGCTCGCTGCAAGACCTTCTCTGCTCGGTC 179</td>
</tr>
<tr>
<td>99049M_W_NA</td>
<td>TGGGACAGCTCCCGACAGCGGCTGCGCCGCTCGCTGCAAGACCTTCTCTGCTCGGTC 179</td>
</tr>
<tr>
<td>99064M_W_NA</td>
<td>TGGGACAGCTCCCGACAGCGGCTGCGCCGCTCGCTGCAAGACCTTCTCTGCTCGGTC 179</td>
</tr>
</tbody>
</table>
```
<table>
<thead>
<tr>
<th>Code</th>
<th>Sequence</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>10107N_W_EU</td>
<td>CAGACAACACC 729</td>
<td></td>
</tr>
<tr>
<td>10106N_W_EU</td>
<td>CAGACAACACC 729</td>
<td></td>
</tr>
<tr>
<td>10152N_W_EU</td>
<td>CAGACAACACC 729</td>
<td></td>
</tr>
<tr>
<td>07019MB</td>
<td>--------------</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 2.4 Alignment of β-tubulin sequences from *M. nivale* and *M. majus*. Primer-binding sites are indicated by shading.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Alignment</th>
<th>Primer-binding sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>10099M_W_EU</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>10098M_W_EU</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>10149M_W_EU</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>99027M_W_EU</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>99049M_W_EU</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>99061M_W_EU</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>99084M_W_EU</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>10107N_W_EU</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>10106N_W_EU</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>99006N_W_EU</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>99069N_W_EU</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>10152N_W_EU</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>96107N_T_NA</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>96103N_T_NA</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>96101N_T_NA</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>10101N_T_EU</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>10082N_T_EU</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>10083N_T_EU</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>07019_OUT</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>07020_OUT</td>
<td>AGGCTTCGCCAACACGATAGCTTCTCTGCGCCCTCTGCGAGGAGCTGCGGGTGACCA</td>
<td>60</td>
</tr>
<tr>
<td>10099M_W_EU</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>10098M_W_EU</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>10149M_W_EU</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>99027M_W_NA</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>99049M_W_NA</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>99061M_W_NA</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>99084N_W_NA</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>10107N_W_EU</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>10106N_W_EU</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>99006N_W_NA</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>99069N_W_NA</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>10152N_W_EU</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>96107N_T_NA</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>96103N_T_NA</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>96101N_T_NA</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>10101N_T_EU</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>10082N_T_EU</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>10083N_T_EU</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>07019_OUT</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>07020_OUT</td>
<td>TGGATGCCGTCCGTGCTGGTCCCTTCGGCCAGCTGTTCCGCCCCGACAACTTCGTCTTCG</td>
<td>120</td>
</tr>
<tr>
<td>10099M_W_EU</td>
<td>GTCAGTCCGGTGCTGGCAACAACTGGGCGAAGGGTCACTACACTGAGGGTGCCGAGCTTG</td>
<td>180</td>
</tr>
<tr>
<td>10098M_W_EU</td>
<td>GTCAGTCCGGTGCTGGCAACAACTGGGCGAAGGGTCACTACACTGAGGGTGCCGAGCTTG</td>
<td>180</td>
</tr>
<tr>
<td>10149M_W_EU</td>
<td>GTCAGTCCGGTGCTGGCAACAACTGGGCGAAGGGTCACTACACTGAGGGTGCCGAGCTTG</td>
<td>180</td>
</tr>
<tr>
<td>99027M_W_NA</td>
<td>GTCAGTCCGGTGCTGGCAACAACTGGGCGAAGGGTCACTACACTGAGGGTGCCGAGCTTG</td>
<td>180</td>
</tr>
<tr>
<td>99049M_W_NA</td>
<td>GTCAGTCCGGTGCTGGCAACAACTGGGCGAAGGGTCACTACACTGAGGGTGCCGAGCTTG</td>
<td>180</td>
</tr>
<tr>
<td>99061M_W_NA</td>
<td>GTCAGTCCGGTGCTGGCAACAACTGGGCGAAGGGTCACTACACTGAGGGTGCCGAGCTTG</td>
<td>180</td>
</tr>
<tr>
<td>99084N_W_NA</td>
<td>GTCAGTCCGGTGCTGGCAACAACTGGGCGAAGGGTCACTACACTGAGGGTGCCGAGCTTG</td>
<td>180</td>
</tr>
<tr>
<td>10107N_W_EU</td>
<td>GTCAGTCCGGTGCTGGCAACAACTGGGCGAAGGGTCACTACACTGAGGGTGCCGAGCTTG</td>
<td>180</td>
</tr>
<tr>
<td>10106N_W_EU</td>
<td>GTCAGTCCGGTGCTGGCAACAACTGGGCGAAGGGTCACTACACTGAGGGTGCCGAGCTTG</td>
<td>180</td>
</tr>
<tr>
<td>99006N_W_NA</td>
<td>GTCAGTCCGGTGCTGGCAACAACTGGGCGAAGGGTCACTACACTGAGGGTGCCGAGCTTG</td>
<td>180</td>
</tr>
<tr>
<td>99069N_W_NA</td>
<td>GTCAGTCCGGTGCTGGCAACAACTGGGCGAAGGGTCACTACACTGAGGGTGCCGAGCTTG</td>
<td>180</td>
</tr>
<tr>
<td>10152N_W_EU</td>
<td>GTCAGTCCGGTGCTGGCAACAACTGGGCGAAGGGTCACTACACTGAGGGTGCCGAGCTTG</td>
<td>180</td>
</tr>
</tbody>
</table>

Appendix 2.5 Alignment of EF-1α sequences from *M. nivale* and *M. majus*. Primer binding sitse are indicated by shading.
Appendix 2.6 Alignment of ITS sequences from *M. nivale* and *M. majus*. Primer binding
sites and *RsaI* restriction sites indicated by shading.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>9007M_W_NA</td>
<td>AGAAATAAGTTAAAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCA 178</td>
</tr>
<tr>
<td>9008N_W_NA</td>
<td>AGAAATAAGTTAAAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCA 178</td>
</tr>
<tr>
<td>10152N_W_EU</td>
<td>AGAAATAAGTTAAAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCA 178</td>
</tr>
<tr>
<td>99007N_W_NA</td>
<td>AGAAATAAGTTAAAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCA 178</td>
</tr>
<tr>
<td>99061M_W_NA</td>
<td>AGAAATAAGTTAAAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCA 178</td>
</tr>
<tr>
<td>99064M_W_NA</td>
<td>AGAAATAAGTTAAAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCA 178</td>
</tr>
<tr>
<td>07020_MB_OUT</td>
<td>TGCTGCCACCGGTGGACTAC</td>
</tr>
<tr>
<td>99019_MB</td>
<td>CTCCAAACCATGTGAACTTACCACTGTTGCCTCGGTGGATGGTGCTGTCTCTCGGGACG</td>
</tr>
<tr>
<td>10096M_W_EU</td>
<td>CTCCAAACCATGTGAACTTACCACTGTTGCCTCGGTGGATGGTGCTGTCTCTCGGGACG</td>
</tr>
<tr>
<td>10151M_W_EU</td>
<td>CTCCAAACCATGTGAACTTACCACTGTTGCCTCGGTGGATGGTGCTGTCTCTCGGGACG</td>
</tr>
<tr>
<td>99069N_W_NA</td>
<td>CTCCAAACCATGTGAACTTACCACTGTTGCCTCGGTGGATGGTGCTGTCTCTCGGGACG</td>
</tr>
<tr>
<td>96107N_T_NA</td>
<td>CTCCAAACCATGTGAACTTACCACTGTTGCCTCGGTGGATGGTGCTGTCTCTCGGGACG</td>
</tr>
<tr>
<td>10082N_T_EU</td>
<td>CTCCAAACCATGTGAACTTACCACTGTTGCCTCGGTGGATGGTGCTGTCTCTCGGGACG</td>
</tr>
<tr>
<td>10083N_T_EU</td>
<td>CTCCAAACCATGTGAACTTACCACTGTTGCCTCGGTGGATGGTGCTGTCTCTCGGGACG</td>
</tr>
<tr>
<td>96103N_T_NA</td>
<td>CTCCAAACCATGTGAACTTACCACTGTTGCCTCGGTGGATGGTGCTGTCTCTCGGGACG</td>
</tr>
<tr>
<td>10101N_T_EU</td>
<td>CTCCAAACCATGTGAACTTACCACTGTTGCCTCGGTGGATGGTGCTGTCTCTCGGGACG</td>
</tr>
<tr>
<td>10102N_T_EU</td>
<td>CTCCAAACCATGTGAACTTACCACTGTTGCCTCGGTGGATGGTGCTGTCTCTCGGGACG</td>
</tr>
<tr>
<td>10152N_W_EU</td>
<td>CTCCAAACCATGTGAACTTACCACTGTTGCCTCGGTGGATGGTGCTGTCTCTCGGGACG</td>
</tr>
<tr>
<td>99061M_W_NA</td>
<td>CTCCAAACCATGTGAACTTACCACTGTTGCCTCGGTGGATGGTGCTGTCTCTCGGGACG</td>
</tr>
<tr>
<td>99049M_W_NA</td>
<td>CTCCAAACCATGTGAACTTACCACTGTTGCCTCGGTGGATGGTGCTGTCTCTCGGGACG</td>
</tr>
<tr>
<td>99064M_W_NA</td>
<td>CTCCAAACCATGTGAACTTACCACTGTTGCCTCGGTGGATGGTGCTGTCTCTCGGGACG</td>
</tr>
<tr>
<td>99027M_W_NA</td>
<td>CTCCAAACCATGTGAACTTACCACTGTTGCCTCGGTGGATGGTGCTGTCTCTCGGGACG</td>
</tr>
<tr>
<td>RsaI sites:</td>
<td>G-TAC-</td>
</tr>
<tr>
<td>9007M_W_NA</td>
<td>TACCCACCGCGTGGACTACGAAAACCTCT-GTTAATTTTTGCAAC-TCTGAACTAAACTA 118</td>
</tr>
<tr>
<td>90064M_W_NA</td>
<td>TACCCACCGCGTGGACTACGAAAACCTCT-GTTAATTTTTGCAAC-TCTGAACTAAACTA 118</td>
</tr>
<tr>
<td>90049M_W_NA</td>
<td>TACCCACCGCGTGGACTACGAAAACCTCT-GTTAATTTTTGCAAC-TCTGAACTAAACTA 118</td>
</tr>
<tr>
<td>99061M_W_NA</td>
<td>TACCCACCGCGTGGACTACGAAAACCTCT-GTTAATTTTTGCAAC-TCTGAACTAAACTA 118</td>
</tr>
<tr>
<td>99007N_W_NA</td>
<td>TACCCACCGCGTGGACTACGAAAACCTCT-GTTAATTTTTGCAAC-TCTGAACTAAACTA 118</td>
</tr>
<tr>
<td>10152N_W_EU</td>
<td>TACCCACCGCGTGGACTACGAAAACCTCT-GTTAATTTTTGCAAC-TCTGAACTAAACTA 118</td>
</tr>
<tr>
<td>99069N_W_NA</td>
<td>TACCCACCGCGTGGACTACGAAAACCTCT-GTTAATTTTTGCAAC-TCTGAACTAAACTA 118</td>
</tr>
<tr>
<td>99064M_W_NA</td>
<td>TACCCACCGCGTGGACTACGAAAACCTCT-GTTAATTTTTGCAAC-TCTGAACTAAACTA 118</td>
</tr>
<tr>
<td>99027M_W_NA</td>
<td>TACCCACCGCGTGGACTACGAAAACCTCT-GTTAATTTTTGCAAC-TCTGAACTAAACTA 118</td>
</tr>
<tr>
<td>99049M_W_NA</td>
<td>TACCCACCGCGTGGACTACGAAAACCTCT-GTTAATTTTTGCAAC-TCTGAACTAAACTA 118</td>
</tr>
<tr>
<td>99064M_W_NA</td>
<td>TACCCACCGCGTGGACTACGAAAACCTCT-GTTAATTTTTGCAAC-TCTGAACTAAACTA 118</td>
</tr>
<tr>
<td>99027M_W_NA</td>
<td>TACCCACCGCGTGGACTACGAAAACCTCT-GTTAATTTTTGCAAC-TCTGAACTAAACTA 118</td>
</tr>
<tr>
<td>99007N_W_NA</td>
<td>TACCCACCGCGTGGACTACGAAAACCTCT-GTTAATTTTTGCAAC-TCTGAACTAAACTA 118</td>
</tr>
<tr>
<td>10152N_W_EU</td>
<td>TACCCACCGCGTGGACTACGAAAACCTCT-GTTAATTTTTGCAAC-TCTGAACTAAACTA 118</td>
</tr>
</tbody>
</table>

* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
10101N_T_EU       AATACCCGCTGAACCTTAAGCATA 497
96103N_T_NA       AATACCCGCTGAACCTTAAGCATA 497
10085N_T_NA       AATACCCGCTGAACCTTAAGCATA 497
99084N_W_NA       AATACCCGCTGAACCTTAAGCATA 497
96107N_T_NA       AATACCCGCTGAACCTTAAGCATA 497
10083N_T_EU       AATACCCGCTGAACCTTAAGCATA 497
10082N_T_EU       AATACCCGCTGAACCTTAAGCATA 497
96101N_T_NA       AATACCCGCTGAACCTTAAGCATA 497
99069N_W_NA       AATACCCGCTGAACCTTAAGCATA 497
10151M_W_EU       AATACCCGCTGAACCTTAAGCATA 497
10096M_W_EU       AATACCCGCTGAACCTTAAGCATA 497
10099M_W_EU       AATACCCGCTGAACCTTAAGCATA 497
10149M_W_EU       AATACCCGCTGAACCTTAAGCATA 497
07019_MB_OUT      AATACCCGCTGAACCTTAAGCATA 495
07020_MB_OUT      AATACCCGCTGAACCTTAAGCATA 496

***************
Chapter 3 Comparative Genomics

3.1 Introduction

3.1.1 General overview of whole-genome analyses

As the availability of DNA and RNA sequencing data has increased rapidly in the past 30 years (GenBank 2012), researchers have developed new tools to allow them to search, compare, analyse, and manipulate increasingly complex data sets. Bioinformatics is a field that uses the tools of computer science to address these challenges in biology (Attwood et al. 2011). One use of bioinformatics is the assembly and analysis of whole-genome sequencing data (Pop and Salzberg 2008).

The first eukaryotic genomes became available in the late 20th century (Goffeau et al. 1996), and were produced with Sanger sequencing (Sanger et al. 1977). However, sequencing technology has advanced rapidly, and a number of different sequencing strategies have become commercially available within the last decade (Shendure and Ji 2008). These techniques have reduced both the cost and the length of time required to produce large amounts of sequencing data (Lister et al. 2009). Genome sequences of a number of organisms have been produced using these technologies in recent years, particularly in species where reference genomes are available (e.g. (Hillier et al. 2008; Ossowski et al. 2008)).

Fungi present an interesting and tractable challenge for de novo whole genome sequencing using next generation methods (Haridas et al. 2011). The de novo sequencing of fungal genomes has already been accomplished using next-generation sequencing alone (Nowrousian et al. 2010). Filamentous ascomycetes are haploid in their vegetative state (Alexopoulos et al. 1996), and the conidia and ascospores of most ascomycota are usually uninucleate (Alexopoulos et al. 1996). These factors facilitate the collection of genetically
homogeneous material and the sequencing of single ascomycetous isolates. Furthermore, the genomes of most filamentous ascomycetes are smaller than 37 Mb (Gregory et al. 2007), which facilitates both the collection of sufficient data for successful assembly and renders the assembly of these sequence fragments computationally feasible.

3.1.2. Sequencing platforms

Until recently, most sequencing was performed using the Sanger method, which, when first developed, amplified a sequence of interest in the presence of both normal deoxynucleic acids (dNTPs) and radioactively-labelled chain-terminating dideoxynucleic acids (ddNTPs). This process produced a pool of oligonucleotides of various lengths that were separated on agarose gels and visualized using x-ray film (Sanger et al. 1977). Sanger sequencing was updated through the use of fluorescently-labelled dNTPs, eliminating the need for radioactivity, and facilitating the automation of detection (e.g. (Ansorge et al. 1986)). Sanger sequencing produces reads that are approximately 1 kb in length (Shendure and Ji 2008).

Although Sanger sequencing is still common, especially for the sequencing of short DNA molecules (such as plasmids and PCR products), newer sequencing technologies have become common in large-scale sequencing projects (e.g. (Li et al. 2009b)). Like Sanger sequencing, Illumina-Solexa sequencing is a sequencing-by-synthesis method, but it has the additional benefit of producing results in real-time (Metzker 2010). The process of Illumina-Solexa sequencing (also abbreviated as SBS, for sequencing-by-synthesis) is illustrated in Figure 3.1. In SBS, as in other next-generation sequencing methods, genomic DNA is first sheared into random fragments (Nielsen et al. 2011). Both ends of the fragments are ligated to adapter segments of DNA, which in turn are complementary to short nucleotide sequences that are attached to a solid
surface, effectively binding the strand to be sequenced to the solid surface (Metzker 2010). Prior to sequencing, the strands are amplified by PCR: because the adapter sequences are complementary to the sequences bound to the solid surface, the adapter segments act as short, double-stranded fragments for DNA polymerase (Figure 3.1). When the amplification is complete, the result is "clusters" of identical sequence fragments. To sequence the fragment, the four fluorescently-labelled dNTPs are introduced into the reaction mixture. Because the dNTPs have been chemically modified to block the 3' hydroxyl group, only a single nucleotide at a time can be introduced (Mardis 2008). When the reaction mixture is washed away, the identity of the nucleotide that has been added is read, the blocking group is cleaved away, and the sequencing procedure is repeated (Mardis 2008). The length of the reads produced by SBS has increased sharply, from 36 bp in 2008 (Shendure and Ji) to 150 bp in 2012 (Quail et al.).

Several other next-generation sequencing platforms also exist. Pyrosequencing, a technique employed by 454 sequencing, uses the pyrophosphate molecule released during the incorporation of a nucleotide onto a growing strand of DNA to generate a "flash" of light using the enzyme luciferase (Mardis 2008). The order of the sequence in question is determined by introducing a single dNTP species at a time into the reaction mixture, and simply recording those which induce a flash of light (Metzker 2010). In 2005, the read length produced by this platform was already 100-150 bp, and as of 2012 it is 700 bp (Liu et al. 2012).

Ion Torrent sequencing (also known as ion semiconductor sequencing) shares some similarities with 454 sequencing. Whereas 454 sequencing is dependent upon the detection of pyrophosphate when a nucleotide is introduced into a growing chain, Ion Torrent sequencing detects the release of a proton (i.e. a hydrogen ion). When a proton is released, the pH of the reaction mixture decreases slightly. This change in acidity is thus detected only when the correct
dNTP solution is introduced onto the reaction surface. As of 2012, the read length of ion torrent sequencing was approximately 200 bp (Quail et al. 2012).

In sequencing by ligation (SOLiD sequencing), as in SBS, the DNA is first cleaved into small fragments and ligated to adapter sequences. A primer sequence that is complementary to the adapter is introduced. Next, fluorescently-labeled priming sequences are introduced into the reaction mixture. These eight-base-pair long (octameric) priming sequences consist of two nucleotides, followed by a set of five nucleotides ligated to a fluorophore. If the first two nucleotides are complementary to the strand to be sequenced, they will bind, and the fluorophore will be detected (Mardis 2008). The octamer is cleaved after the fifth base, and the next octamer is introduced. In this way, the nucleotides at the 1st and 2nd, the 6th and 7th, 11th and 12th positions, and so on, will be determined. The procedure is then repeated, but rather than using a priming sequence that is the full length of the adapter sequence, a primer that is one bp shorter is used. In this way, the first octamer will reveal the 0th (adapter) and the 1st bases, and subsequent rounds will determine identities of the 5th and 6th and the 10th and 11th positions, and so on. This entire process is repeated until the full sequence of the fragment has been determined (Mardis 2008), and the maximum read length in 2012 was 50 bp (Quail et al.).

Although all four of these four next-generation methods are currently in use (in addition to several other methods not discussed here), each one offers distinct advantages and disadvantages in terms of the overall cost, speed, and accuracy of the sequencing results. Due to the short read lengths of all next-generation sequencing technologies relative to Sanger sequencing, one disadvantage shared by all of these methods is the difficulty in re-assembling repetitive regions (Alkan et al. 2011). Because each nucleotide is queried twice, SOLiD sequencing has a high level of accuracy relative to other next-generation methods (Metzker
2010). However, this same benefit also increases both the duration and the cost of a sequencing run performed using SOLiD sequencing (Metzker 2010). Pyrosequencing offers long read lengths relative to the other next-generation methods (up to 1 kb as of 2013 (Roche Diagnostics Corporation 2013)), but the necessary reagents are expensive and the method has a higher error rate for single-nucleotide repeats ("homopolymers"; (Metzker 2010)). Ion torrent sequencing generates data extremely quickly, but otherwise shares the same problems as pyrosequencing and has a read length comparable to other next-generation methods (Quail et al. 2012). By comparison, a single run of SBS is slower, but less expensive than many of the other next-generation sequencing options (Metzker 2010; Quail et al. 2012). Sequencing by synthesis generates a large number of reads in each run, permitting deep coverage for genomic sequencing, which facilitates later assembly (Quail et al. 2012). For these reasons, SBS has been successfully used in the de novo sequencing of organisms such as *Ailuropoda melanoleuca* (the giant panda) (Li et al. 2009b) and *Homo sapiens* (Simpson et al. 2009).

### 3.1.3 Genome assembly and protein prediction

The short length of the reads produced by next-generation sequencing techniques creates a challenge for genome assembly similar to the way that a puzzle with more pieces is more difficult to assemble than one with fewer pieces. The use of paired-end sequencing attempts to mitigate this issue. Paired-end sequencing involves sequencing a fragment of known length from both ends (Fullwood et al. 2009). Because the approximate full length of the DNA fragment is known, as well as the length of the two sequenced regions, these sequences can be re-assembled more easily because the distance between them is known (Figure 3.2).
In addition to the advantage offered by paired-end sequencing, most freely-available sequencing programs, including SOAPdenovo (Li et al. 2009c; Li et al. 2008), Velvet (Zerbino and Birney 2008), and ABySS (Simpson et al. 2009), use de Brujin graphs to associate short groups of aligned reads (Figure 3.3). Before preparing a de Brujin graph, the raw, short reads are aligned (Zerbino and Birney 2008). The alignments comprise the nodes of the de Brujin graph, and the nodes are then associated with one another based on their similarities (i.e. potential for overlap) (Zerbino and Birney 2008). In this way, small "pockets" of similarity are associated, and a longer sequence is thus assembled based on the alignments of the short reads (Zerbino and Birney 2008).

Although the goals of any individual study may vary, the prediction of protein-coding sequences is often an important prerequisite to whole-genome comparison. The identification of protein-coding sequences may facilitate a variety of comparisons, including the identification of novel genes that may be related to the lifestyle of an organism of interest (O’Connell et al. 2012). The identification of gene families that contain duplications or deletions may also help to clarify evolutionary relationships (Chang and Duda 2012). In addition, the examination of an organism's protein-coding sequences may reveal the genetic mechanisms responsible for an organism's unique abilities, such as cool-temperature survival (Methé et al. 2005) or pathogenicity (O’Connell et al. 2012).

Although the basic concept of identifying protein-coding regions within the genome may appear simple on the surface, the problem of identifying only the functional protein sequences from among the millions of base pairs in a sequenced genome is not trivial. For example, the frequency (and even the presence) of introns may vary between taxa (Stajich et al. 2006), and the
presence of pseudogenes (truncated or otherwise frame-shifted former protein-coding sequences) (Nelson and Cox 2004) may confound algorithms designed to detect protein-coding sequences.

Many of the gene-finding programs developed to date are based on hidden Markov models (e.g. (Borodovsky and Lukashin 1998; Stanke et al. 2004)). A Markov chain is a statistical model that describes the probability of a system assuming any particular state in the future, given its current state; the probability that the system will assume any of the possible state identities in the future is dependent only on the current state of the system (Stamp 2012). In a hidden Markov model (HMM), a Markov process is occurring wherein the identity of the state is hidden, but the outcome resulting from this state is observable (Stamp 2012). For example, in the context of gene discovery, the relative frequencies of the four nucleotides in coding regions of DNA may differ from that in non-coding regions, and the identity of these hidden states (i.e. as coding vs. non-coding) for a given segment of DNA can be elucidated through the observed pattern of nucleotides and the known probabilities of transitioning between each nucleotide within either state and between the two states themselves (Eddy 2004). A more accurate model to predict the state of a given sequence as coding or non-coding would incorporate more than one observation (i.e. nucleotide); this is the basis of a generalized hidden Markov model (GHMM), wherein a single (hidden) state produces multiple observations (Stanke 2003). The gene-finding program AUGUSTUS is based on this model and uses a string of nucleotides to estimate the current state of the string as one of a multitude of possible states (e.g. an intron, an intron-exon boundary, promoter region, etc.) (Stanke 2003). Outside of pure de novo prediction methods, gene predictions may also be based on the sequencing information available for related species or, in some cases, previously-sequenced members of the same species (Sleator 2010); AUGUSTUS also uses the coding sequences of related organisms to guide its predictions.
3.1.4 Whole-genome comparisons

As genome sequencing technology has become more widely accessible, the genomes of several hundred fungal species (Choi et al. 2013) have been published. Some of these fungal species include pathogens of plants and animals, and some researchers have sequenced two or more representatives of related groups to search for trends in the genomic origin of pathogenicity and in the overall arrangement of these fungal genomes (e.g. (Amselem et al. 2011; Gao et al. 2011; Jackson et al. 2009; Schirawski et al. 2010; Sharpton et al. 2009)). Among several sequenced pairs of species, including the insect pathogens *Metarhizium anisopliae* and *M. acridium* (Gao et al. 2011), the human pathogens *Candida albicans* and *C. dubliniensis* (Jackson et al. 2009), the plant pathogens *Colletotrichum graminicola* and *C. higginsianum* (O’Connell et al. 2012), *Sclerotinia sclerotiorum* and *Botrytis cinerea* (Amselem et al. 2011), and the maize pathogens *Ustilago maydis* and *Sporisorium reilianum*, high degrees of synteny were reported between each pair. Similarly, when three species of *Aspergillus* were sequenced, synteny in both coding and non-coding regions of the genome was reported between all species (Galagan et al. 2005).

In each of their comparisons between the genomes of pathogenic fungal species, the researchers investigated pathogenic differences and the genetic basis of host specificity among the species studied by identifying predicted genes with putative pathogenic functions. For example, although both are entomopathogens, *M. acridium* is a locust-specific pathogen whereas *M. anisopliae* attacks a variety of insect species. Despite this difference, the genomes shared nearly 90% amino acid homology among their predicted genes (Gao et al. 2011). Both also possessed several predicted proteins that were homologous to reported plant pathogenesis-related
proteins (such as hydrophobins), in addition to proteases that are predicted to play a role in the degradation of the cuticle of insects. There were no apparent differences in the relative amounts of types of degradative enzymes present in the genome that could readily explain the observed host preferences. In contrast, the maize pathogens *Ustilago maydis* and *Sporisorium relianum* possessed dissimilar effector proteins, which may reflect their differences in tissue type preference and infection symptoms (Schirawski et al. 2010). The hemibiotrophic pathogens *Colletotrichum higginsianum* and *C. graminicola* also possessed differences in their predicted proteins related to their functions as pathogens (O’Connell et al. 2012). The genome of *C. higginsianum* contained more than twice as many pectinases as the genome of *C. graminicola*; in contrast, the genome of *C. graminicola* encoded a more diverse family of cellulases than those found in *C. higginsianum*. These differences correlate with each species' preferred host: whereas *C. higginsianum* is primarily a pathogen of dicots, which possess cell walls that are richer in pectin, *C. graminicola* is a pathogen of grasses and cereals (monocots), which have cell walls that are rich in cellulose.

In the necrotrophic plant pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea* (Amselem et al. 2011), both genomes shared homologs of pathogenesis-related genes, such as genes involved in programmed cell death in plants. Similarly, both possessed more genes related to oxidative phosphorylation than other ascomycetes, which the authors suggested may be related to their production of oxalic acid during the infection process. However, both genomes carried fewer copies of pectinases, cellulases, and hemicellulases than most other plant pathogenic fungi surveyed, such as *Magnaporthe oryzae* and *Giberella zeae*.

Trends in the presence and absence of pathogenicity-related genes were also detected when the genomes of the human fungal pathogens *Coccidioides immitis* and *C posadasii* were
compared to those of the non-pathogenic but closely-related *Uncinocarpus reesii* and the pathogenic but more distantly-related *Histoplasma capsulatum* (Sharpton et al. 2009). All four species belong to the Onygenales and were compared to the genomes of previously-sequenced members of the Sordariomycetes, including several species of *Aspergillus* and *Penicillium*, which are both members of the Eurotiales. In these comparisons, all members of the Onygenales, including the phytopathogenic species, were either lacking entirely or possessed far fewer predicted proteins homologous to protein families that are directly related to plant degradation (e.g. cellulases, cutinases, pectin lyases) relative to the Eurotiales studied. However, the genomes of the pathogenic *Coccidioides* species studied possessed a large number of serine proteases and keratinases relative to the other species examined, which may be related to their role as animal pathogens, rather than soil saprobes.

Whole-genome sequencing has also identified trends in the physical arrangement of the genomes of pathogenic strains relative to non-pathogenic strains. As early as 1992, the Fot1 transposable element, detected in *Fusarium oxysporum*, was proposed as an important source of genetic variation (Daboussi et al.). Transposable elements (TEs) are fragments of DNA that can change their position within a genome (Hartl and Clark 2007) (chapter 4). There are two general classes of transposable elements, denoted as class I and class II. Class I elements are also called retrotransposons because the TE sequence that was previously integrated into the genome is transcribed into RNA before being reverse-translated back into DNA to be re-integrated into the genome at a different position (Daboussi and Capy 2003). Class II elements move throughout the genome by being excised as DNA and simply re-locating to a new position (Daboussi and Capy 2003). Members of both class I and class II TEs have been identified in filamentous Ascomycetes (Daboussi and Capy 2003) and transposon-rich regions are physically associated
with genes that confer host specificity in several Ascomycete plant pathogens, including *Magnaporthe oryzae* (Dean et al. 2005; Thon et al. 2006), *Alternaria alternaria* (Hatta et al. 2002), and *Verticillium* spp. (Amyotte et al. 2012).

Furthermore, large-scale genomic differences have been reported between filamentous fungi with differing plant hosts. When the genomes of *Fusarium oxysporum* f. sp. *lycopersici*, *F. verticillioides*, and *F. graminearum* were sequenced, *F. oxysporum* was found to contain 19 Mb of sequences that were not shared with the other *Fusarium* species examined (Ma et al. 2010). These sequences contained 74% of the TE sequences identified in this species, and although some of these sequences were found on chromosomes that were shared with the other species, the majority were organized into four unique chromosomes. In addition to containing the majority of TEs, the unique segments of the *F. oxysporum* genome also contained a high concentration of predicted genes that shared a high level of sequence identity with effectors, virulence factors, and proteins that are involved in signal transduction (i.e. genes that may play a role in pathogenicity). When the genome of the *F. oxysporum* f. sp. *lycopersici* isolate sequenced was compared to those of different strains of *F. oxysporum* (one a pathogen of *Arabidopsis* and one belonging to f. sp. *vasinfectum*), the genes within the unique region were not shared with the other strains examined, suggesting that these genes may play a key role in host-specific pathogenicity. Taken together, this evidence suggests that the comparison of closely-related genomes may permit the identification of expanded gene families that may help to explain some of the unique characteristics of the pathogenicity of a species of interest.
3.1.5 Objectives

The primary objective of this project was to obtain and assemble whole-genome sequences for isolates of *M. nivale* and *M. majus*. These genome sequences were compared to each other and to the whole-genome sequences of other filamentous ascomycetes to determine the amount of variation within and between different species of *Microdochium*. The whole-genome information was used to design a single primer set that amplifies both species but distinguishes between them by producing amplicons of different sizes.

3.2 Materials and Methods

3.2.1 DNA extraction, quantification, and sequencing

Mycelium for DNA extraction was prepared as described in Chapter 2. The genomic DNA for NGS was extracted using either the Qiagen DNeasy Plant Mini Kit (Qiagen, Mississauga, Canada) (*M. majus* isolate 99049 and *M. nivale* isolate 11037) or the PowerSoil DNA isolation kit (Mo Bio, Carlsbad, CA, USA) (*M. nivale* isolates 12262 and 10106, *M. majus* isolate 10095, and *M. bolleyi* isolate 07020). The *M. bolleyi* DNA was extracted by Fang Shi (Hsiang lab) using the Qiagen method. For both methods, a total of 200 mg of fungal tissue was processed in two separate 100 mg batches for each of the isolates sequenced. The Qiagen extraction was performed according to the manufacturer's instructions with the following modifications. During the tissue homogenization step, an initial volume of 200 µL of buffer AP1 was added to a tube containing 100 mg of fungal tissue and approximately 50 mg of sterile, acid-washed sea sand (Fisher, Fair Lawn NJ, USA). The mycelium was then homogenized as described in Chapter 2. An additional 200 µL of buffer AP1 was added to the tube before proceeding with the remaining steps described in the manufacturer's protocol. During the final
elution step, DNA was eluted with two washes of 50 µL rather than a single 100 µL rinse. The PowerSoil extraction was performed according to the manufacturer's instructions following the "alternative lysis method" with the following modifications. During each of the three vortex steps described in the alternative lysis protocol, the samples were vortexed for 20 s, rather than 3-4 s. The final elution was performed using 50 µL of buffer per sample, and the moistened columns were incubated at room temperature for 10 minutes before the final centrifugation. Following extraction, all DNA was stored at -20 °C. The quality and quantity of the DNA was assessed by electrophoresis through agarose gels as described in Chapter 2.

The quantity of DNA sent for sequencing and the sequencing facility utilized for each reaction is described in Table 3.1. For M. majus isolate 99049, an Illumina Genome Analyser IIx platform was used specifying a single full lane of 75 bp paired-end sequencing. For all other genomes sequenced, an Illumina HiSeq 2000 platform was used specifying 100 bp paired-end sequencing from 1/4 to 1/8 of single lanes (multiplexed).

3.2.2 Genome assembly and gene prediction

Genome assemblies and analysis were conducted using either SHARCNET (Shared Hierarchical Academic Research Computing NETwork; www.sharcnet.ca), WestGrid (www.westgrid.ca), or local UBUNTU 12.01 servers with 16 gb (HP16) or 32 gb (GIGA32) of RAM. Where applicable, sequencing data obtained in BAM format were converted to FASTQ using the program SAMtools (Li et al. 2009a) prior to further analysis. All of the genomes studied were assembled using the programs ABySS v. 1.3.4 (Simpson et al. 2009) and SOAPdenovo v 1.05 (Li et al. 2008) and GapCloser v. 1.12 (Li et al. 2008) with odd-numbered kmers between 29 and 63. Assembly quality was assessed by examining the N50 value and by
examining the total number of scaffolds produced by the program. The N50 value is a statistic that provides an estimate of assembly quality. This value is calculated by sorting all contigs (or scaffolds) by size and then identifying the contig size at which 50% of all bases in the assembly are included in contigs that are larger and smaller than this value (Haridas et al. 2011). A higher N50 value reflects larger contigs or scaffolds. The word contig is an abbreviation of contiguous, and refers to the sequences that are produced from overlapping reads in a sequence assembly (Haridas et al. 2011). By comparison, scaffolds are composed of neighbouring contigs that are a known distance apart based on paired-read information, but that may be joined by unknown bases due to insufficient sequencing coverage of the adjoining regions (Haridas et al. 2011).

For *M. majus* isolate 99049, two separate rounds of assemblies were performed. The first assembly was performed using the methods described above, using the raw sequencing reads. In the second round of assemblies, the contigs from the highest N50 assembly of the first round were used as single-end inputs and assembled again. Predicted genes were prepared for the assemblies of each genome with highest N50 using AUGUSTUS v.2.5.5 (Stanke et al. 2004). The predicted gene set of *Magnaporthe grisea* (included in the program) was used to train the algorithm. The predicted genes were annotated using the web interface of FastAnnotator (available at http://fastannotator.cgu.edu.tw; (Chen et al. 2012)). Additional detail was added to the annotations by comparing the putative protein family (Pfam) accession numbers (Punta et al. 2012) assigned to the predicted genes to the full Pfam database downloaded from http://pfam.sanger.ac.uk. The script "annotate_genes.pl" was used to format the annotated sequences. An example of the scripts and configuration files used for assembly, gene prediction, and annotation are found in Appendix 3.2, Appendix 3.3, Appendix 3.4, Appendix 3.5, and Appendix 3.6.
3.2.3 Whole-genome comparisons and identification of unique genes

BLAST databases for the predicted gene sets and the assembled scaffold sequences were prepared using the command "makeblastBLASTdb" included with the standalone BLAST package BLAST+ v.2.2.25 (Altschul 1990). To identify predicted genes that are unique to *M. nivale* and/or *M. majus*, the predicted gene sets of the sequenced *Microdochium* genomes were searched against each other and against a set of fifteen other Sordariomycete genomes, including six members of the Xylariales (Table 3.7) using the tBLASTn algorithm. The resulting data were parsed (parse_m9.pl; Appendix 3.7) to remove comments. For each of the *Microdochium* genomes studied, the results were then summarized in a matrix (make_simple_table_v2.pl, Appendix 3.8) wherein the presence or absence of each predicted gene in all of the other genomes and/or genome categories (the *Microdochium*, Xylariales, and other Sordariomycetes) was recorded as either present (1) or absent (0). This script also applied a filter to exclude predicted genes that were shorter than 100 amino acids (or 300 bp) in length and to impose a maximum e-value score. For these comparisons, three different maximum e-value thresholds were used: 1e-05, 1e-20, and 1e-50. Finally, this table was queried (summarize_with_files_v2.pl, Appendix 3.9) to identify genes that were unique to the given genome, or that were only found within certain categories of interest (e.g. only among *M. nivale* isolates from wheat).

To assess the relationships within *Microdochium*, the full gene sequences (including introns) of ten predicted protein sequences that were apparently unique to *Microdochium* (based on the methods described above) were collected from the genomes of all of the *Microdochium* isolates sequenced. These sequences were concatenated and aligned using ClustalX as described in Section 2.2.5, and maximum likelihood, neighbour-joining, and maximum parsimony trees were constructed using PAUP* v. 4.0 to visualize the relationships between the isolates.
sequenced. For each tree, 100 bootstrap replicates were calculated. The identities of these genes are listed in Table 3.10 and their alignment is found in Appendix 3.15.

In addition to these comparisons, the synteny between the whole genomes of *M. majus* isolate 99049 and *M. nivale* isolates 11037 and 12262 was investigated using the program Mauve v. 2.3.1 (Darling et al. 2004). Prior to this analysis, each of the *Microdochium* genomes was independently aligned in Mauve to the genome of *Magnaporthe grisea* (isolate 70-15, assembly version 6), downloaded from the Broad Institute, [http://www.broadinstitute.org/](http://www.broadinstitute.org/). These "pre-aligned" genomes were then re-aligned together (without *M. grisea*).

During the multiple-gene comparisons performed in Chapter 2, the *M. bolleyi* EF-1α sequence could not be amplified using a combination of the primers EFNivF, EFMicF, and EFMicR. To investigate this hypothesis, a putative EF-1α sequence was identified in the *M. bolleyi* predicted gene set by using the *M. nivale* and *M. majus* partial EF-1α sequences amplified in Chapter 2 to query the *M. bolleyi* 07020 predicted gene set using the BLASTx algorithm. The EFNivF and EFMicF primer sequences, as well as the reverse complement of the EFMicR primer sequence were aligned with the putative *M. bolleyi* EF-1α nucleotide sequence using ClustalX.

### 3.2.4 Design of species-specific primers

The nucleotide sequences of the predicted genes of *M. nivale* isolate 11037 were used to query a BLAST database constructed from the nucleotide sequences of the *M. majus* 99049 genome. The output of this search was parsed using a script (find_genes_of_diff_length.pl, Appendix 3.10) to identify genes that had both a maximum e-value of 1E-05 and a minimum length difference of 50 bp. The full sequences of genes that met these criteria were extracted
from the genome (including both introns and exons), and were aligned using ClustalW. The alignments were sorted manually to assess both the quality of the match and the presence or absence of total length differences. Primers were designed (Section 2.2.6) for candidates which possessed both a) highly conserved regions that were at least 300 bp apart (for primer design purposes), and b) a sequence length difference of at least 50 bp between the highly conserved regions. These primers were synthesized by Laboratory Services Division, University of Guelph (Guelph, Canada) and were tested according to the protocol for the ITS PCR (Section 2.2.4), with the exception of the annealing temperature. A list of the primers tested and their annealing temperatures are found in Table 3.5. Each of the primers was tested with at least two *M. majus* isolates (one each from North America and from Europe), and at least two *M. nivale* isolates (one each from wheat and from turf).

### 3.2.5 Identification of putative pathogen-related genes

A list of potential pathogen-related genes were identified in the six sequenced *Microdochium* genomes using a list of pathogen-host interaction (PHI) genes obtained from PHI-base (Winneburg et al. 2006). These sequences were queried against BLAST databases constructed from the predicted protein sequences using the BLASTp algorithm. The presence or absence of these sequences among the predicted proteins was then assessed using the script "compare_phi_results.pl" (Appendix 3.11). The number of matches for each gene was assessed, and all matches with a length of at least 100 bp and an e-value of 1e-25 or lower were tabulated. Where possible, the function of genes with a large difference in the number of matches among the *Microdochium* genomes were noted (Table 3.9).
3.2.6 Identification of putative transposable elements

Where available, the protein sequences of transposable elements previously identified in other filamentous ascomycetes was downloaded from GenBank. Nucleotide sequences were downloaded for some sequences when protein sequences were unavailable. These sequences were selected to represent the major families of Class I and Class II elements that have been reported from filamentous ascomycetes. A list of the sequences used is found in Table 3.11. These sequences were queried against the BLAST databases that had been constructed from the scaffold genome assemblies of all six Microdochium sp. isolates examined using the tBLASTx or tBLASTn algorithm as appropriate. All matches with e-values of less than 1e-05 were collected for further analysis. To avoid counting the same putative TEs as hits to more than one of the query sequences, the hits in the genome were only counted as putative TEs if they were not within 500 bp of another hit. The positions of these putative TEs relative to the putative PHI genes identified as described above (Section 3.2.6) were compared using the script "check_proximity.pl" (Appendix 3.12), and all putative PHI genes that were within 5 kb of a putative TE were noted. Duplicate matches of similar TE-PHI pairs on the same scaffold were eliminated using the script "strip_duplicates.pl" (Appendix 3.13).

3.3 Results

3.3.1 Genome sequencing, assembly, and protein prediction

The whole-genome sequences of M. majus isolates 99049 and 10095, M. nivale isolates 11037, 12262, and 10106, and M. bolleyi isolate 07020 were obtained as raw reads in either FASTQ (99049, 10095, 12262 10106, 07020) or BAM (11037) format. The total number of raw reads obtained for each genome is summarized in Table 3.3.
The genome assembly programs Velvet, SOAPdenovo, and ABySS were used to assemble the genome of *M. majus* isolate 99049. Of these three programs, Velvet yielded the poorest results (Table 3.4) based on the N50 values and the total genome size. This program was not used to assemble the other genomes, which were obtained later. For all genomes, kmers 25-63 were successfully assembled by both SOAPdenovo and ABySS. The genome sizes of both *M. majus* isolates were 35.9 Mb. For *M. nivale* turf isolate 11037, the draft genome sequence was 37.0 Mb in length; for the wheat isolates 12262 and 10106, the genome sizes were 36.7 and 37.1 Mb, respectively. The genome of *M. bolleyi* isolate 07020 was 38.2 Mb in length.

The program AUGUSTUS was used to predict protein sequences from all of the genomes studied, and the total number of proteins predicted from the best assemblies of each of the genomes are summarized in Table 3.3. For each of the genomes, only the assembly with the largest N50 was used for protein prediction. All of the *M. nivale* and *M. majus* isolates sequenced, with the exception of *M. nivale* turf isolate 11037, had between 11.3 to 11.7x10^3 predicted proteins. The *M. nivale* turf isolate had 12x10^3 predicted proteins, and the *M. bolleyi* isolate had 13x10^3 predicted proteins. The number of predicted genes from each genome that were successfully annotated using FastAnnotate is noted in Table 3.3.

### 3.3.2 Whole-genome comparisons

Several comparisons were made between and among the whole-genome data of the six *Microdochium* genomes sequenced in these experiments. These data were also compared to whole-genome data from seven species of the Xylariales and thirteen other non-Xylariales Sordariomycetes. The results of these comparisons, with e-value maxima of 1e-05, 1e-20, and 1e-50, are found in Table 3.5, Table 3.6, and Table 3.7, respectively.
At an e-value maximum of 1e-05, the two *M. majus* genomes, 10095 and 99049, shared 99.8 and 99.7%, respectively, of all genes. When the e-value minimum was decreased to 1e-50, these species shared 89.4 and 92.7% of all genes. For the genomes of *M. nivale* isolates 12262 and 10106, which were both isolated from wheat, 97.9 and 98.0% of genes were shared at an e-value minimum of 1e-05. At an e-value of 1e-50, 92.9 and 89.4 % of all genes, respectively, were shared between these isolates. The *M. nivale* turf isolate 11037 shared 92.2% of its genes with both of the *M. nivale* wheat isolates at an e-value minimum of 1e-05. At an e-value minimum of 1e-50, this isolate shared 78.5% of its genes with the two *M. nivale* wheat isolates. Between *M. nivale* and *M. majus*, at an e-value of 1e-05, isolate 11037 shared 91.6% of its genes with both *M. majus* genomes; the *M. nivale* wheat isolates 12262 and 10106 shared 95.5% and 92.9% of their genes with the *M. majus* isolates, respectively. The *M. majus* isolates 99049 and 10095 shared 92.5 and 89.9% of their genes with 11037. Isolate 99049 shared 95.3% with both 12262 and 10106, and 10095 shared 93.2% with both *M. nivale* wheat isolates. At an e-value minimum of 1e-50, isolate 11037 shared 78.0% of its genes with both *M. majus* isolates. The wheat isolates 12262 and 10106 shared 81.8 and 78.9% of their genes with the two *M. majus* isolates. The *M. majus* isolates 99049 and 10095 shared 79.7 and 77.1% of their genes, respectively, with turf isolate 11037. Isolate 99049 shared 82.6% and 10095 shared 79.7% of its genes with the two *M. nivale* wheat isolates.

At an e-value minimum of 1e-05, the *M. bolleyi* isolate shared 81.0% of its genes with *M. nivale* turf isolate 11037, 80.3% with both of the *M. nivale* wheat isolates, and 80.5% of its genes with the *M. majus* isolates. At an e-value minimum of 1e-50, the *M. bolleyi* isolate shared 67.7% of its genes with *M. nivale* turf isolate 11037, 66.8% of its genes with both the *M. nivale* wheat isolates, and the *M. majus* isolates.
In addition to these comparisons, ten predicted genes that were not found in any of the Sordariomycete genomes examined but that were found in all of the *Microdochium* genomes were concatenated and were used to construct maximum likelihood (ML), neighbour-joining (NJ), and maximum parsimony (MP) trees, each with 100 bootstrap replicates (Figure 3.4). In all three trees, the two *M. majus* isolates formed a single clade with 100% bootstrap support. Similarly, the two *M. nivale* isolates from wheat also formed a single clade with 100% bootstrap support in all trees. The *M. nivale* turf isolate 11037 did not group with the other *M. nivale* isolates in any of the trees. In both the NJ and the ML trees, the isolate grouped with the *M. majus* isolates (with 61 and 87% bootstrap support, respectively). In the MP tree, the isolate formed a tricotomy with the *M. bolleyi* isolate and the clade containing the remaining isolates. In all three trees, the *M. bolleyi* isolate formed either a dichotomy (NJ and ML) or a trichotomy (MP) with the other *Microdochium* isolates included in the analysis. In all three trees, the relative branch lengths linking the two *M. majus* isolates were shorter than the branch lengths for the *M. nivale* wheat isolates.

A comparison of the whole-genome synteny among the *Microdochium* genomes was attempted using the program Mauve; however, the inability to assemble the *Microdochium* spp. genomes into chromosomes prevented the extraction of meaningful results from this comparison. Despite pre-aligning the genomes with the genome of *M. grisea*, the synteny of the *Microdochium* genomes could not be assessed.

### 3.3.3 Development of species-specific primers

Of the 228 predicted gene sequences aligned, 14 were selected as good candidates for further investigation based on their possession of two conserved regions (appropriate for primer
design) that were separated by at least 300 bp, as well as sufficient indels to yield a length difference of at least 50 bp in between these two conserved regions. Primers were designed and ordered for seven of these genes. The results of the primer tests are summarized in Table 3.5. Of the primer sets tested, only Med5 and 371 yielded results that differed consistently between the two species of interest.

### 3.3.4 Identification of pathogenesis-related genes

A total of 2,614 pathogen-host interaction (PHI) genes (Winnenburg et al. 2006) were searched against the predicted protein sequences of all of the *Microdochium* genomes. A total of 1,831 genes were shared among all of the *Microdochium* predicted gene sets examined. Of the 2,614 genes included in the analysis, 451 were not identified in any of the predicted gene sets.

In *M. bolleyi* isolate 07020, 1,873 of the PHI genes were found, including nine that were not shared with any of the other species examined; conversely, the other five predicted gene sets shared a total of eight genes that were found in all predicted gene sets except *M. bolleyi*. In *M. nivale* turf isolate 11037, 1,880 PHI genes were found, including five that were not shared with any of the other species examined. In *M. nivale* wheat isolates 12262 and 10106, 1,870 and 1,866 PHI genes, respectively, were identified. One gene was identified in isolate 12262 that was not shared with any of the other *Microdochium* predicted gene sets; there were no unique genes in isolate 10106. There were four PHI genes that were found in both of the *M. nivale* wheat isolates that were not shared with any of the other isolates examined. In *M. majus* isolates 99049 and 10095, 1,866 and 1,871 PHI genes were identified, respectively. Two genes that were not shared with any of the other sequenced isolates were identified in isolate 99049, but none were observed in isolate 10095. Among the *M. majus* isolates, there were two genes that were found in both
predicted gene sets that were not found in any of the other *Microdochium* predicted gene sets. Among the 451 PHI genes that were not found in any of the *Microdochium* predicted gene sets, 148 of these were bacterial genes, 19 were from Oomycetes, and 87 were from Basidiomycetes; the remaining "missing" genes were all from members of the Ascomycota.

Because the *Microdochium* species examined shared a nearly-identical set of putative PHI genes, the relative number of hits for each of the PHI genes was also assessed to determine whether copy number differences could be detected. For 1,736 of the PHI genes, the number of putative copies (within the restrictions imposed on the search) found in all six of the *Microdochium* genomes was identical. For the remaining PHI genes, differences were observed in the number of putative matches found in the predicted gene sets. The number of matches in the *M. nivale* and *M. majus* genomes was nearly identical for the majority of PHI genes identified. The total number of matches in each genome differed by fewer than ten for all of the PHI genes assessed, with the exception of a single PHI gene which had 28 copies in both the turf isolate *M. nivale* 11037 and in *M. bolleyi* 07020 and only seven matches in all four of the other genomes. In addition, there were 47 genes that had at least ten more matches in the *M. bolleyi* genome than in any of the other genomes, and 39 genes that had ten or fewer matches in the *M. bolleyi* genome relative to any of the others. The function of many of the PHI genes that yielded these discrepancies were unknown; however, Table 3.9 lists PHI genes with differing copy numbers in the *Microdochium* genomes for which putative functions were available.

### 3.3.5 Identification of putative transposable elements and comparison to PHI genes

A total of 24 transposable element sequences representing the major groups of TEs found in filamentous ascomycetes were downloaded from GenBank, and were used to query the
Microdochium spp. genomes. Of the 24 TE query sequences, four did not have a match in any of the Microdochium genomes. These four sequences included the Class II MITE-like sequences guest from N. crassa and mimp from F. oxysporum, the mutator-like sequence from N. parvum, and the Class I gypsy mars integrase from A. immersus. Both of the M. majus isolates examined possessed matches for the same 17 of 24 TE sequences queried. Among the M. nivale isolates examined, each isolate displayed a slightly different pattern of which TEs were found. The turf isolate 11037 possessed matches to only 13 of the 24 TE sequences, whereas the wheat isolates 12262 and 10106 matched 15 and 16 isolates, respectively. One TE sequence, the Class I skippy gag sequence from F. oxysporum, possessed at least one match in all of the genomes except the two M. nivale wheat isolates. The M. bolleyi isolate 07020 possessed at least one match to 19 of the 24 TE sequences, and one TE had a match in M. bolleyi only. Within these categories, the total number of putative TEs identified in each species is listed in Table 3.12. When the positions of the putative TE sequences were compared to those of the PHI genes, over 50% of the putative TEs were found to be within 5kb of a PHI gene.

3.3.6 Identification of a putative EF-1α sequence in the genome of M. bolleyi

When the M. nivale and M. majus partial EF-1α nucleotide sequences obtained in Chapter 2 were queried against the M. bolleyi 07020 predicted protein set, the top match for all of the sequences was predicted protein g6339. When the primers that had been used to amplify the partial EF-1α sequences from M. nivale and M. majus were aligned with the nucleotide sequence of this predicted gene, the reverse complement of the reverse primer was found to be a perfect match to the predicted gene sequence; however, both of the forward primer sequences had several mismatches with the predicted gene sequence. More importantly, the predicted binding
site of the reverse primer was upstream relative to the predicted binding sites of both forward primers. The alignment of these sequences is found in Appendix 3.14.

3.4 Discussion

Next-generation sequencing technology was used to obtain de novo genome sequences for two isolates of *M. majus*, three isolates of *M. nivale*, and one isolate of *M. bolleyi*. The isolates were sent for sequencing at separate times, with the genome of 99049 obtained first, followed by 11037, then 12262, and finally 10096, 10106, and 07020 all at the same time. The raw sequencing reads were successfully assembled into scaffolds, which were used to predict protein sequences from all isolates. The assembled genomes and predicted proteins were then used to perform comparisons between these sequenced isolates.

For *M. majus* isolate 99049 only, two separate rounds of genome assembly were conducted. In the first round, the raw sequencing reads were used as the input for the assembly programs, whereas in the second assembly round, the contigs obtained from the first assembly were used as the input. This second assembly significantly increased the size of the scaffolds obtained, from an N50 of 4,106 to 96,968. This large increase in contig size increased the utility of the genome data by increasing the number of predicted genes on each contig. This was particularly useful in MAT gene identification and synteny studies (Chapter 5).

For the remaining genome sequences, periodic re-assembly of the genomes using the same methodology and/or using newer versions of the assembly programs as they were released yielded only modest improvements in the overall assembly quality based on the N50 score. Two factors are likely to have contributed to the poor quality of the initial assembly of 99049. First, the genome sequencing data for 99049 was obtained approximately six months earlier than any
of the other data, and was generated using a read length of only 75 bp with an insert size of 200 bp, whereas all of the other assemblies had a read length of 100 bp with an insert size of at least 300 bp. This difference reflects the rapid improvement of sequencing technology and methodology within this short time frame. Second, while the sequencing of 99049 was conducted in Vancouver, British Columbia, the remaining isolates were sequenced in Montréal, Québec. The relative proximity of the Montréal sequencing facility decreased the time that the samples spent in transit, thus minimizing the risk of unfavourable conditions such as warm temperatures, which may have decreased the overall quality of the DNA of 99049. The comparatively poor quality of the sequencing data from 99049 relative to the other M. nivale and M. majus sequencing data, in addition to similar results obtained for other fungal genomes sequenced at these two facilities within the Hsiang lab (personal communication), led to the use of the Québec sequencing facility for the other genomes sequenced after 99049.

Aside from the difficulties discussed in assembling the genome of 99049, the other Microdochium genomes were readily assembled using SOAPdenovo and ABySS. The N50 values for the final assemblies used in subsequent comparisons ranged between 96 kb (99049) to 371 kb (10106) for the M. nivale and M. majus data. The N50 obtained for the M. bolleyi genome was over 2 Mb in length; this is over five times greater than that obtained for any of the other Microdochium genomes. This difference could not be readily explained; the approximate quantity and concentration of the M. bolleyi DNA was similar to that of the other Microdochium DNA sequenced at the same time, and the quality of this assembly also exceeded that of any of the other filamentous Ascomycetes sequenced in the Hsiang lab (unpublished results), including genomes that have been obtained after the M. bolleyi data. This unexpected result is the subject of active research.
A nearly-identical number of predicted genes were obtained from both of the *M. majus* isolates sequenced, despite the fact that the second genome obtained (10095) had over 50% greater sequencing coverage than the 99049 genome. A similar trend was observed between the two *M. nivale* wheat genomes, 12262 and 10106, where the genome of 12262 had over 150% greater sequencing coverage than 10106, yet the 10106 genome yielded slightly more predicted genes. In addition, a similar number of predicted genes were obtained for all of the *M. nivale* and *M. majus* genomes sequenced. By comparison, there were nearly 1,000 more predicted genes in *M. bolleyi* relative to the *M. nivale* isolate with the highest number of predicted genes, whereas within the *M. nivale* and *M. majus* genomes, the difference in number between the isolate with the highest (11037) and the lowest (99049) number of predicted genes was just over 700. These similarities between and among the *M. nivale* and *M. majus* predicted gene sets suggest that, although there are likely protein-coding sequences that were not detected by AUGUSTUS (e.g. sequences that may have been interrupted by scaffold boundaries, or that received poor sequencing coverage simply due to random chance), the number of protein-coding genes among these species was similar. In addition, the fully sequenced and annotated genomes of over 300 fungi are available (Choi et al. 2013), including several phytopathogenic ascomycetes. Among these pathogenic ascomycetes, the number of predicted genes range between 5,854 (for *Blumeria graminis*) to 17,735 (in *Fusarium oxysporum*) (Raffaele and Kamoun 2012). The number of predicted genes for all of the *Microdochium* genomes sequenced in these experiments is well within this range. Together, these results suggest that the predicted gene sets obtained for the genomes sequenced were likely reasonably complete.

Using the genomes of *M. nivale* and *M. majus*, several sets of primers were designed with the goal of developing a method to rapidly distinguish between *M. nivale* and *M. majus* by
amplifying a single band of a different size for each species. A total of seven primer sets were designed and tested towards this goal. Despite using the genomic data of *M. majus* isolate 99049 and *M. nivale* isolate 11037, these primers proved difficult to design and optimize. In most cases, the primers failed to yield results that differed in a consistent manner between the two species. Instead, multiple bands were often produced, which for most primers rendered the results for the two species indistinguishable. This discrepancy may have been caused because the primers were intentionally designed to span highly diverse regions (i.e. introns), and these regions may have differed in length between individuals, rather than between species as was originally intended.

To assess the variation both within and between these species, reciprocal genome-vs-genome tBLASTn searches (predicted genes vs scaffold sequences) were conducted. The results of these comparisons were assessed at three different e-value stringencies, with the goal of assessing the similarities and differences among and between the predicted genes of these species. At the highest stringency level (i.e. the smallest e-value), a smaller overall number of shared genes was expected, but the proportion of genes that were shared between the more closely-related species (e.g. within the *M. majus* genomes) was expected to grow. At the least stringent e-values, the overall number of genes shared between the genomes studied was expected to grow. When these comparisons were conducted, the differences between the three stringency levels were smaller than expected. The same general trends were observed between all three stringency levels.

The tBLASTn algorithm was chosen for this analysis for two primary reasons. First, we were interested in identifying homologs of putative protein-coding sequences, rather than assessing the overall similarity of the genome, including non-coding regions. Second, by choosing tBLASTn, the predicted gene sets were compared to the whole-genome nucleotide
sequences of each sequenced isolate. This facilitated the identification of protein-like sequences that were not included in the predicted protein set; this exclusion could occur for sequences that are truncated or otherwise non-functional homologs of sequences that may have functional analogs in the other isolates, or could simply be the result of an "oversight" by AUGUSTUS.

Among the *M. nivale* and *M. majus* isolates examined, an unexpected trend was observed regarding the dissimilarity of the *M. nivale* turf isolate 11037 to the two *M. nivale* wheat isolates also sequenced. Despite originating from North America and Europe, respectively, the two *M. nivale* wheat isolates 12262 and 10106 shared 97.9 and 95.6%, respectively, of their predicted protein sequences with each other at an e-value cutoff of 1e-05. By comparison, each of these isolates shared an average of 95.5 and 92.9%, respectively, of their predicted genes with the two sequenced *M. majus* isolates, and 93.1 and 90.1%, respectively, of their genes with *M. nivale* turf isolate 11037. Similarly, at an e-value cutoff of 1e-05, the turf isolate shared only slightly more of its predicted genes with the two *M. nivale* wheat isolates (92.2% average) compared to the two *M. majus* isolates (91.6% average). Very broadly, this isolate was thus found to be approximately equally dissimilar to *M. nivale* from wheat and to *M. majus* (but was still more similar to either of these groups than to *M. bolleyi*). This trend is tentatively in agreement with earlier assertions that *M. nivale* possesses a relatively high level of intra-specific variation and that there may be distinct sub-groups within this population (Lees et al. 1995). Similarly, because the other four *M. nivale* and *M. majus* strains examined were all originally isolated from wheat, this shared host plant origin may at least partially explain their relative similarity.

To explore this observation further and to visualize the relationships between the sequenced *Microdochium* genomes, the concatenated sequences of ten predicted proteins that were found in the genomes of all *Microdochium* species surveyed, but which were also not found
among any of the non-Microdochium genomes, were used to construct bootstrapped neighbour-joining, maximum likelihood, and maximum parsimony trees. All three trees were rooted with *M. bolleyi*, and yielded the same broad conclusions: both the *M. nivale* wheat isolates and the *M. majus* isolates each formed single clades with 100% bootstrap support. Surprisingly, in both the NJ and ML trees, the *M. nivale* turf isolate 11037 grouped with the node containing the *M. majus* isolates with strong bootstrap support. In the MP tree, the *M. nivale* 11037 isolate formed a trichotomy with *M. bolleyi* and with the clade containing the remaining isolates.

Although the relative dissimilarity between the sequenced turf isolate of *M. nivale* and the wheat isolates is an unexpected an interesting observation, it is not possible to draw broad conclusions regarding the overall similarity (or lack thereof) between all *M. nivale* turf isolates compared to those from wheat based on the observations of only five genomes (only one of which was from turf). Instead, this finding demonstrates the importance of future work to investigate this dissimilarity. At least one European and one additional North American *M. nivale* turf isolate should be included in this analysis before attempting to form a more general hypothesis regarding the apparent importance of host origin in *M. nivale*.

In both the whole-genome vs. genome BLAST comparison and the Microdochium-specific gene tree, *M. bolleyi* was found to be more dissimilar to the other Microdochium spp. genomes than any of the other genomes were to one another. This was easily rationalized, as *M. nivale* and *M. majus* were considered to be a single species until 2005 (Glynn et al.). In addition, *M. bolleyi* is primarily associated with the roots of graminaceous species (Murray and Gadd 1981), whereas *M. majus* and *M. nivale* attack the stem and grain of grasses and cereals (Parry 1990).
Differences between *M. bolleyi* and the other *Microdochium* genomes examined were also detected among the number of matches in the predicted gene sets of these genomes against the genes from the pathogen-host interaction (PHI) database. Although the identities of the PHI genes found among these genomes were similar, the relative number of matches among the genomes was more variable, especially between *M. bolleyi* relative to the other *Microdochium* genomes. A total of 86 PHI genes had at least ten more or at least ten fewer matches in *M. bolleyi* than in any other *Microdochium* genome examined. Although the putative functions of the majority of these genes were not available based on their GenBank summaries, their inclusion in the PHI database clearly suggests that they may play a role in the pathogenicity of these species. One possible explanation for this difference is that *M. bolleyi's* grows on or inside roots (Murray and Gadd 1981), rather than in above-ground plant parts as do *M. nivale* and *M. majus* (Parry 1990). In addition, *M. bolleyi* is described as a "weak" root pathogen (Murray and Gadd 1981), and has been found in non-damaging interactions with plant roots (Domsch and Gams 1972).

Putative transposable elements from both classes were identified in the draft genome sequences of all of the *Microdochium* spp. isolates examined. Representatives of nine out of the ten TE superfamilies were identified in the *Microdochium* genomes, with the exception of the Class II MITE-like elements. In addition to identifying TEs, the positions of these sequences relative to the PHI genes were identified to investigate the hypothesis that pathogenicity-related genes are often associated with TEs (e.g. (Amyotte et al. 2012; Hatta et al. 2002)). Of the putative TEs identified in the *Microdochium* genomes, more than 50% of the sequences were found within 5 kb of a PHI gene, suggesting that *Microdochium* follows this general trend.
This chapter describes the generation of whole-genome sequencing data for a total of six Microdochium strains, consisting of two isolates of *M. majus*, two isolates of *M. nivale* from wheat, one isolate of *M. nivale* from turfgrass, and one isolate of *M. bolleyi*. All of these *Microdochium* genomes were found to share the majority of their predicted gene sets, including the subset of predicted genes that were predicted to play a role in pathogenicity. Despite originating from different continents, the two *M. majus* genomes examined were found to share over 99% of their predicted genes. A similar trend was observed for the two *M. nivale* wheat isolates, but, surprisingly, the one *M. nivale* turf isolate examined was approximately equally dissimilar to the *M. nivale* wheat and the *M. majus* genomes. This difference was supported by both the total number of predicted genes shared between the genome and by building phylogenetic trees based on the alignments of genes that were found only among the *Microdochium* genomes. However, the relative number of matches to genes in a database of pathogen-related genes was found to be nearly identical between all of the *M. nivale* and *M. majus*, whereas this similarity was not shared with the predicted genes of *M. bolleyi*. The whole-genome information prepared in this chapter was used to inform all future research including, most significantly, the search for mating-type genes in these species (Chapter 5).
3.5 References for Chapter 3


142


Shim, W.B., and Dunkle, L.D. 2003. CZK3, a MAP kinase kinase kinase homolog in Cercospora zeae-maydis, regulates cercosporin biosynthesis, fungal development, and pathogenesis. Molecular Plant Microbe Interactions 16(9): 760-768.


Table 3.1 Summary of DNA quantity and sequencing facility utilized for genome sequencing of six *Microdochium* isolates.

<table>
<thead>
<tr>
<th>Species</th>
<th>Isolate Number</th>
<th>Quantity of DNA submitted (µg)</th>
<th>Insert size (bp)</th>
<th>Read length (bp)</th>
<th>Sequencing Facility</th>
<th>Location of Sequencing Facility</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>M. majus</em></td>
<td>99049</td>
<td>1</td>
<td>200</td>
<td>75</td>
<td>BC Genomics Sequencing Centre</td>
<td>Vancouver, BC</td>
</tr>
<tr>
<td></td>
<td>10095</td>
<td>4</td>
<td>400</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>M. nivale</em></td>
<td>11037</td>
<td>3</td>
<td>300</td>
<td>100</td>
<td>Génome Québec Innovation Centre</td>
<td>Montréal, QC</td>
</tr>
<tr>
<td></td>
<td>12262</td>
<td>2</td>
<td>400</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10106</td>
<td>2</td>
<td>400</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>M. bolleyi</em></td>
<td>07020</td>
<td>4</td>
<td>400</td>
<td>100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3.2 Sordariomycetes genomes included in whole-genome comparisons against *Microdochium* spp.

<table>
<thead>
<tr>
<th>Order</th>
<th>Family</th>
<th>Genus</th>
<th>Species</th>
<th>Genome Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glomerellales</td>
<td>Plectosphaerellaceae</td>
<td><em>Acremonium</em></td>
<td><em>alcalophilum</em></td>
<td>JGI*</td>
</tr>
<tr>
<td></td>
<td>Glomerellaceae</td>
<td><em>Colletotrichum</em></td>
<td><em>graminearum</em></td>
<td>Hsiang lab</td>
</tr>
<tr>
<td>Hypocreales</td>
<td>mitosporic</td>
<td><em>Verticillium</em></td>
<td><em>dahlieae</em></td>
<td>JGI</td>
</tr>
<tr>
<td></td>
<td>Hypocreales</td>
<td></td>
<td><em>albo-atrum</em></td>
<td>Broad Institute</td>
</tr>
<tr>
<td>Magnaporthales</td>
<td>Magnaporthaceae</td>
<td><em>Magnaporthea</em></td>
<td><em>oryzae</em></td>
<td>Broad Institute</td>
</tr>
<tr>
<td></td>
<td>Sordariaceae</td>
<td><em>Neurospora</em></td>
<td><em>crassa</em></td>
<td>Broad Institute</td>
</tr>
<tr>
<td>Sordariales</td>
<td>Chaetomiaceae</td>
<td><em>Sporotrichum</em></td>
<td><em>thermophile</em></td>
<td>JGI</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Thielavia</em></td>
<td><em>terrestris</em></td>
<td>JGI</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xylariales</td>
<td>Xylariaceae</td>
<td><em>Annulohypoxylon</em></td>
<td><em>stygium</em></td>
<td>Hsiang lab</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Daldinia</em></td>
<td><em>eschscholtzii</em></td>
<td>Hsiang lab</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Hypoxylon</em></td>
<td><em>sp. (CO)</em></td>
<td>JGI</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>sp. (EC)</em></td>
<td>JGI</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Pestalotiopsis</em></td>
<td><em>nigrescens</em></td>
<td>Hsiang lab</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>theae</em></td>
<td></td>
</tr>
</tbody>
</table>

* JGI = Joint Genome Institute
Table 3.3 Summary of genome assembly, protein prediction, and predicted gene annotation statistics for sequenced *Microdochium* genomes

<table>
<thead>
<tr>
<th>Species</th>
<th>Isolate</th>
<th>Reads</th>
<th>Highest N50 (bp)</th>
<th>Genome size (bp)</th>
<th>Coverage</th>
<th>Number of predicted genes</th>
<th>N50 of predicted genes</th>
<th>Number of predicted genes annotated</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>M. majus</em></td>
<td>99049</td>
<td>24,527,354</td>
<td>96,968</td>
<td>35,892,675</td>
<td>137</td>
<td>11,343</td>
<td>1,962</td>
<td>7,617</td>
</tr>
<tr>
<td></td>
<td>10095</td>
<td>77,722,046</td>
<td>260,023</td>
<td>35,930,000</td>
<td>216</td>
<td>11,651</td>
<td>1,986</td>
<td>7,667</td>
</tr>
<tr>
<td><em>M. nivale</em></td>
<td>11037</td>
<td>52,250,482</td>
<td>139,319</td>
<td>37,010,607</td>
<td>282</td>
<td>12,060</td>
<td>1,956</td>
<td>8,080</td>
</tr>
<tr>
<td></td>
<td>12262</td>
<td>56,244,760</td>
<td>309,358</td>
<td>36,683,753</td>
<td>307</td>
<td>11,430</td>
<td>1,989</td>
<td>7,666</td>
</tr>
<tr>
<td></td>
<td>10106</td>
<td>44,321,898</td>
<td>371,751</td>
<td>37,085,241</td>
<td>119</td>
<td>11,714</td>
<td>1,992</td>
<td>7,669</td>
</tr>
<tr>
<td><em>M. bolleyi</em></td>
<td>07020</td>
<td>41,555,928</td>
<td>2,035,661</td>
<td>38,158,599</td>
<td>109</td>
<td>13,047</td>
<td>2,016</td>
<td>8,060</td>
</tr>
</tbody>
</table>
Table 3.4 Genome assembly statistics for *M. majus* assembly with Velvet, SOAPdenovo, and ABysS for odd-numbered kmers 29-59.

Note that gap-closing was not performed for this comparison.

<table>
<thead>
<tr>
<th>kmer</th>
<th>SOAPdenovo</th>
<th></th>
<th>ABysS</th>
<th></th>
<th>Velvet</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of Scaffolds</td>
<td>Genome size (Mb)</td>
<td>N50</td>
<td>Number of Scaffolds</td>
<td>Genome size (Mb)</td>
</tr>
<tr>
<td>29</td>
<td>5,113</td>
<td>34.4</td>
<td>2,057</td>
<td>12,004</td>
<td>36.2</td>
</tr>
<tr>
<td>31</td>
<td>2,838</td>
<td>34.6</td>
<td>2,893</td>
<td>10,500</td>
<td>36.2</td>
</tr>
<tr>
<td>33</td>
<td>1,634</td>
<td>34.8</td>
<td>4,218</td>
<td>8,996</td>
<td>36.1</td>
</tr>
<tr>
<td>35</td>
<td>1,119</td>
<td>34.9</td>
<td>6,222</td>
<td>7,948</td>
<td>36.1</td>
</tr>
<tr>
<td>37</td>
<td>811</td>
<td>35.0</td>
<td>8,650</td>
<td>7,337</td>
<td>36.3</td>
</tr>
<tr>
<td>39</td>
<td>717</td>
<td>35.1</td>
<td>10,413</td>
<td>6,814</td>
<td>36.2</td>
</tr>
<tr>
<td>41</td>
<td>678</td>
<td>35.1</td>
<td>11,454</td>
<td>6,283</td>
<td>36.1</td>
</tr>
<tr>
<td>43</td>
<td>684</td>
<td>35.1</td>
<td>12,113</td>
<td>5,827</td>
<td>36.0</td>
</tr>
<tr>
<td>45</td>
<td>603</td>
<td>35.1</td>
<td>11,756</td>
<td>5,321</td>
<td>35.9</td>
</tr>
<tr>
<td>47</td>
<td>621</td>
<td>35.1</td>
<td>11,208</td>
<td>5,024</td>
<td>36.2</td>
</tr>
<tr>
<td>49</td>
<td>656</td>
<td>35.1</td>
<td>10,238</td>
<td>4,666</td>
<td>36.1</td>
</tr>
<tr>
<td>51</td>
<td>720</td>
<td>35.1</td>
<td>9,262</td>
<td>4,378</td>
<td>35.9</td>
</tr>
<tr>
<td>53</td>
<td>811</td>
<td>35.1</td>
<td>8,133</td>
<td>4,078</td>
<td>35.6</td>
</tr>
<tr>
<td>55</td>
<td>985</td>
<td>35.0</td>
<td>6,948</td>
<td>3,808</td>
<td>35.3</td>
</tr>
<tr>
<td>57</td>
<td>1,282</td>
<td>34.9</td>
<td>5,820</td>
<td>3,536</td>
<td>35.8</td>
</tr>
<tr>
<td>59</td>
<td>1,730</td>
<td>34.8</td>
<td>4,731</td>
<td>3,435</td>
<td>35.3</td>
</tr>
</tbody>
</table>

* Assembly at this kmer size did not complete
Table 3.5 Species-specific primers designed and tested with at least two isolates each of *M. nivale* and *M. majus*

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence</th>
<th>Ta (°C)</th>
<th>% GC</th>
<th>Predicted band size (bp)</th>
<th>Observed band size(s) (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>M. nivale</em></td>
<td><em>M. majus</em></td>
</tr>
<tr>
<td>Mm.Mn_impB_2340F</td>
<td>ACCTTGCTGACGGGACC</td>
<td>57.3</td>
<td>65</td>
<td>962</td>
<td>1115</td>
</tr>
<tr>
<td>Mm.Mn_impB3440R</td>
<td>TCGGCAGCAGGACCGCT</td>
<td>55.9</td>
<td>69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mm.Mn_AP2a_3180F</td>
<td>CCGTGTTGCTCAAGCTG</td>
<td>54.9</td>
<td>59</td>
<td>2107</td>
<td>1546</td>
</tr>
<tr>
<td>Mm_Mn_AP2a_4720R</td>
<td>AGGTCTGTGCCAGTCTG</td>
<td>53.5</td>
<td>63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mm.Mn_Med5_3260F</td>
<td>AGACGGACCCCTTTGCTC</td>
<td>57.3</td>
<td>65</td>
<td>628</td>
<td>556</td>
</tr>
<tr>
<td>Mm.Mn_Med5_3800R</td>
<td>AGCAAGTCGGCGCTCG</td>
<td>55.9</td>
<td>69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mm.Mn_GH31_3684F</td>
<td>GCTTCGCCGCACTTCCC</td>
<td>59.8</td>
<td>71</td>
<td>709</td>
<td>800</td>
</tr>
<tr>
<td>Mm_Mn_GH31_4483R</td>
<td>TCCTCAACGTCCGGGCC</td>
<td>59.8</td>
<td>71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mn.Mm_Mn371_3198F</td>
<td>ACGTGCACTGGACCG</td>
<td>55.9</td>
<td>69</td>
<td>603</td>
<td>458</td>
</tr>
<tr>
<td>Mn.Mm_Mn371_3801R</td>
<td>TCAACGGCATSRCTCGCC</td>
<td>58.4</td>
<td>61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mn.Mm_HSP70_1041F</td>
<td>CTCGAACATCCAGGTTGGT</td>
<td>59.5</td>
<td>58</td>
<td>332</td>
<td>472</td>
</tr>
<tr>
<td>Mn.Mm_HSP70_1373R</td>
<td>GCGAGTCGATCTCAGATGG</td>
<td>58.4</td>
<td>61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mn.Mm_C6TF_53F</td>
<td>TGTGGAACACGATCCTCG</td>
<td>56.3</td>
<td>56</td>
<td>1012</td>
<td>880</td>
</tr>
<tr>
<td>Mn.Mm_C6TF_1065sR</td>
<td>GTCTGCAGCCGACTCGA</td>
<td>55.9</td>
<td>69</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* In one isolate only
Table 3.6 Comparisons between the predicted gene sequences from *M. nivale*, *M. majus*, and *M. bolleyi* against each other and against the genomes of 6 (non-*Microdochium*) members of the Xylariales and 9 non-Xylariales members of the Sordariomycetes. Comparisons were performed using tBLASTn with an e-value cutoff of 1e-05.

<table>
<thead>
<tr>
<th>Isolate*</th>
<th>Mn11037</th>
<th>Mn12262</th>
<th>Mn10106</th>
<th>Mm99049</th>
<th>Mm10095</th>
<th>Mb07020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of predicted genes</td>
<td>12,060</td>
<td>11,430</td>
<td>11,714</td>
<td>11,343</td>
<td>11,651</td>
<td>13,047</td>
</tr>
<tr>
<td>Total number of genes &gt; 100 bp</td>
<td>11,792</td>
<td>11,206</td>
<td>11,224</td>
<td>11,083</td>
<td>11,159</td>
<td>12,519</td>
</tr>
<tr>
<td>Number of genes with no hits</td>
<td>397</td>
<td>10</td>
<td>10</td>
<td>16</td>
<td>9</td>
<td>1,543</td>
</tr>
<tr>
<td>Number of genes in all categories</td>
<td>9,551</td>
<td>9,103</td>
<td>9,073</td>
<td>9,110</td>
<td>9,041</td>
<td>9,352</td>
</tr>
<tr>
<td>Number of genes not found in the NM*-NX*-Sordariales</td>
<td>1,378</td>
<td>1,685</td>
<td>1,753</td>
<td>1,583</td>
<td>1,720</td>
<td>220</td>
</tr>
<tr>
<td>Number of genes in NM-Xylariales not also found in the Sordariales</td>
<td>219</td>
<td>188</td>
<td>184</td>
<td>199</td>
<td>189</td>
<td>9,499</td>
</tr>
<tr>
<td>Number of genes in the Sordariales not shared with any <em>Microdochium</em></td>
<td>24</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total number of genes shared with specific genomes:</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mn11037</td>
<td>NA§</td>
<td>10,639</td>
<td>10,563</td>
<td>10,500</td>
<td>10,478</td>
<td>10,568</td>
</tr>
<tr>
<td>Mn12262</td>
<td>11,117</td>
<td>NA</td>
<td>11,197</td>
<td>10,815</td>
<td>10,857</td>
<td>10,482</td>
</tr>
<tr>
<td>Mn10106</td>
<td>11,127</td>
<td>11,193</td>
<td>NA</td>
<td>10,818</td>
<td>10,854</td>
<td>10,481</td>
</tr>
<tr>
<td>Mm99049</td>
<td>11,045</td>
<td>10,918</td>
<td>10,884</td>
<td>NA</td>
<td>11,143</td>
<td>10,499</td>
</tr>
<tr>
<td>Mm10095</td>
<td>11,047</td>
<td>10,915</td>
<td>10,880</td>
<td>11,054</td>
<td>NA</td>
<td>10,505</td>
</tr>
<tr>
<td>Total number of genes found only in wheat Mn</td>
<td>NA</td>
<td>474</td>
<td>560</td>
<td>312</td>
<td>396</td>
<td>NC</td>
</tr>
<tr>
<td>Turf Mn</td>
<td>NA</td>
<td>0</td>
<td>4</td>
<td>27</td>
<td>35</td>
<td>NC</td>
</tr>
<tr>
<td><em>M. bolleyi</em></td>
<td>51</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>All Mn only</td>
<td>94</td>
<td>161</td>
<td>207</td>
<td>1</td>
<td>1</td>
<td>NC</td>
</tr>
<tr>
<td>All Mm only</td>
<td>65</td>
<td>1</td>
<td>3</td>
<td>114</td>
<td>140</td>
<td>NC</td>
</tr>
</tbody>
</table>

150
<table>
<thead>
<tr>
<th>Isolate*</th>
<th>Mn11037</th>
<th>Mn12262</th>
<th>Mn10106</th>
<th>Mn99049</th>
<th>Mn10095</th>
<th>Mb07020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mn AND Mm only</td>
<td>272</td>
<td>669</td>
<td>703</td>
<td>578</td>
<td>717</td>
<td>636</td>
</tr>
</tbody>
</table>

* Mn = M. nivale; Mm = M. majus; Mb = M. bolleyi
† NX = Non-xylariales
‡ NS = Non-Sordariomycete
§ NA = Comparison not applicable
|| NC = Calculation was not performed
Table 3.7 Comparisons between the predicted gene sequences from *M. nivale*, *M. majus*, and *M. bolleyi* against each other and against the genomes of 6 (non-*Microdochium*) members of the Xylariales and 9 non-Xylariales members of the Sordariomycetes. Comparisons were performed using tBLASTn with an e-value cutoff of 1e-20.

<table>
<thead>
<tr>
<th>Isolate*</th>
<th>Mn11037</th>
<th>Mn12262</th>
<th>Mn10106</th>
<th>Mm99049</th>
<th>Mm10095</th>
<th>Mb07020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of predicted genes</td>
<td>12,060</td>
<td>11,430</td>
<td>11,714</td>
<td>11,343</td>
<td>11,651</td>
<td>13,047</td>
</tr>
<tr>
<td>Total number of genes &gt; 100 bp</td>
<td>11,792</td>
<td>11,206</td>
<td>11,224</td>
<td>11,083</td>
<td>11,159</td>
<td>12,519</td>
</tr>
<tr>
<td>Number of genes in all categories</td>
<td>896</td>
<td>53</td>
<td>75</td>
<td>51</td>
<td>49</td>
<td>2,116</td>
</tr>
<tr>
<td>Number of genes in NM-Xylariales not also found in the Sordariales</td>
<td>8,609</td>
<td>8,252</td>
<td>8,241</td>
<td>8,251</td>
<td>8,246</td>
<td>8,467</td>
</tr>
<tr>
<td>Number of genes not found in the NM1-NX1-Sordariales</td>
<td>1,729</td>
<td>2,422</td>
<td>2,434</td>
<td>2,299</td>
<td>2,401</td>
<td>1,332</td>
</tr>
<tr>
<td>Number of genes in NM-Xylariales not also found in the Sordariales</td>
<td>332</td>
<td>310</td>
<td>287</td>
<td>298</td>
<td>293</td>
<td>284</td>
</tr>
<tr>
<td>Number of genes not found in the Sordariales not shared with any <em>Microdochium</em></td>
<td>27</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>8,660</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total number of genes shared with specific genomes:</th>
<th>Mn11037</th>
<th>Mn12262</th>
<th>Mn10106</th>
<th>Mm99049</th>
<th>Mm10095</th>
<th>Mb07020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheat Mn</td>
<td>NA</td>
<td>944</td>
<td>1,049</td>
<td>442</td>
<td>488</td>
<td>NC</td>
</tr>
<tr>
<td>Turf Mn</td>
<td>NA</td>
<td>2</td>
<td>6</td>
<td>50</td>
<td>58</td>
<td>NC</td>
</tr>
<tr>
<td><em>M. bolleyi</em></td>
<td>86</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>only in:</td>
<td>Mn11037</td>
<td>Mn12262</td>
<td>Mn10106</td>
<td>Mn99049</td>
<td>Mn10095</td>
<td>Mb07020</td>
</tr>
<tr>
<td>----------------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>All Mn only</td>
<td>93</td>
<td>533</td>
<td>619</td>
<td>2</td>
<td>2</td>
<td>NC</td>
</tr>
<tr>
<td>All Mm only</td>
<td>65</td>
<td>2</td>
<td>2</td>
<td>405</td>
<td>532</td>
<td>NC</td>
</tr>
<tr>
<td>Mn and Mm only</td>
<td>224</td>
<td>691</td>
<td>670</td>
<td>679</td>
<td>701</td>
<td>806</td>
</tr>
</tbody>
</table>

* Mn = M. nivale; Mm = M. majus; Mb = M. bolleyi
† NX = Non-xylariales
‡ NS = Non-Sordariomycete
§ NA = Comparison not applicable
|| NC = Calculation was not performed
Table 3.8 Comparisons between the predicted gene sequences from *M. nivale*, *M. majus*, and *M. bolleyi* against each other and against the genomes of 6 (non-Microdochium) members of the Xylariales and 9 non-Xylariales members of the Sordariomycetes. Comparisons were performed using tBLASTn with an e-value cutoff of 1e-50.

<table>
<thead>
<tr>
<th>Isolate*</th>
<th>Mn11037</th>
<th>Mn12262</th>
<th>Mn10106</th>
<th>Mm99049</th>
<th>Mm10095</th>
<th>Mb07020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of predicted genes</td>
<td>12,060</td>
<td>11,430</td>
<td>11,714</td>
<td>11,343</td>
<td>11,651</td>
<td>13,047</td>
</tr>
<tr>
<td>Total number of genes &gt; 100 bp</td>
<td>11,792</td>
<td>11,206</td>
<td>11,224</td>
<td>11,083</td>
<td>11,159</td>
<td>12,519</td>
</tr>
<tr>
<td>Number of genes with no hits</td>
<td>1,785</td>
<td>487</td>
<td>641</td>
<td>494</td>
<td>632</td>
<td>3,082</td>
</tr>
<tr>
<td>Number of genes in all categories</td>
<td>6,694</td>
<td>6,531</td>
<td>6,495</td>
<td>6,503</td>
<td>6,533</td>
<td>6,652</td>
</tr>
<tr>
<td>Number of genes not found in the NM*-NX†-Sordariales</td>
<td>2,545</td>
<td>3,529</td>
<td>3,403</td>
<td>3,445</td>
<td>3,329</td>
<td>1,983</td>
</tr>
<tr>
<td>Number of genes in NM-Xylariales not also found in the Sordariales</td>
<td>516</td>
<td>477</td>
<td>462</td>
<td>479</td>
<td>463</td>
<td>482</td>
</tr>
<tr>
<td>Number of genes in the Sordariales not shared with any Microdochium</td>
<td>30</td>
<td>8</td>
<td>12</td>
<td>5</td>
<td>9</td>
<td>6,917</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total number of genes shared with specific genomes:</th>
<th>Mn11037</th>
<th>Mn12262</th>
<th>Mn10106</th>
<th>Mm99049</th>
<th>Mm10095</th>
<th>Wheat Mn</th>
<th>Turf Mn</th>
<th>M. bolleyi</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NA$^*$</td>
<td>9,126</td>
<td>9,054</td>
<td>9,039</td>
<td>8,982</td>
<td>8,829</td>
<td>98</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>9,464</td>
<td>NA</td>
<td>10,463</td>
<td>9,372</td>
<td>9,288</td>
<td>8,715</td>
<td>32</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>9,487</td>
<td>10,622</td>
<td>NA</td>
<td>9,378</td>
<td>9,298</td>
<td>8,711</td>
<td>21</td>
<td>NC$^|$</td>
</tr>
<tr>
<td></td>
<td>9,418</td>
<td>9,354</td>
<td>9,244</td>
<td>NA</td>
<td>10,411</td>
<td>8,718</td>
<td>102</td>
<td>NC$^|$</td>
</tr>
<tr>
<td></td>
<td>9,436</td>
<td>9,363</td>
<td>9,251</td>
<td>10,511</td>
<td>NA</td>
<td>8,735</td>
<td>113</td>
<td>NC$^|$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total number of genes found only in:</th>
<th>Wheat Mn</th>
<th>Turf Mn</th>
<th>M. bolleyi</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NA</td>
<td>96</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>1,405</td>
<td>32</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>1,340</td>
<td>102</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>446</td>
<td>113</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>46</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NC$^|$</td>
<td>NC$^|$</td>
<td>NC$^|$</td>
</tr>
<tr>
<td>Isolate</td>
<td>Mn11037</td>
<td>Mn12262</td>
<td>Mn10106</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>All Mn only</td>
<td>120</td>
<td>1,064</td>
<td>1045</td>
</tr>
<tr>
<td>All Mm only</td>
<td>84</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Mn and Mm only</td>
<td>323</td>
<td>627</td>
<td>588</td>
</tr>
</tbody>
</table>

* Mn = M. nivale; Mm = M. majus; Mb = M. bolleyi
† NX = Non-xylariales
‡ NS = Non-Sordariomycete
§ NA = Comparison not applicable
|| NC = Calculation was not performed
Table 3.9 Fungal pathogen-host interaction (PHI) genes with highly variable copy numbers among *Microdochium* predicted gene sets that may play a role in pathogenicity or fungicide resistance (Mb = *Microdochium bolleyi*; Mm = *M. majus*; Mn = *M. nivale*)

<table>
<thead>
<tr>
<th>Accession Number</th>
<th>Function</th>
<th>Function Reference</th>
<th>Mb 07020</th>
<th>Mm 10095</th>
<th>Mm 99049</th>
<th>Mn 10106</th>
<th>Mn 12262</th>
<th>Mn 11037</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAC49410</td>
<td>Maackiain detoxification</td>
<td>(Convert et al. 1996)</td>
<td>0</td>
<td>35</td>
<td>32</td>
<td>38</td>
<td>30</td>
<td>36</td>
</tr>
<tr>
<td>AAL37947</td>
<td>Osmosensing histidine kinase</td>
<td>(Cui et al. 2002)</td>
<td>4</td>
<td>16</td>
<td>14</td>
<td>15</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>MGG_06847</td>
<td>Transcriptional repressor TUP1</td>
<td>(Elias-Villalobos et al. 2011)</td>
<td>5</td>
<td>16</td>
<td>14</td>
<td>15</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>BAC67162</td>
<td>ATP transporter</td>
<td>(Lee et al. 2005)</td>
<td>15</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>AAP72037</td>
<td>MAP kinase kinase kinase CZK3</td>
<td>(Shim and Dunkle 2003)</td>
<td>7</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>AAS48112</td>
<td>Chimeric spermidine synthase-saccharopine dehydrogenase</td>
<td>(Kingsbury et al. 2004)</td>
<td>2</td>
<td>20</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
</tbody>
</table>
Table 3.10 Accession numbers for ten randomly-selected putative unique predicted gene sequences identified in *Microdochium* spp.

<table>
<thead>
<tr>
<th></th>
<th><em>M. majus</em> 99049</th>
<th><em>M. majus</em> 10095</th>
<th><em>M. nivale</em> 10106</th>
<th><em>M. nivale</em> 11037</th>
<th><em>M. nivale</em> 12262</th>
<th><em>M. bolleyi</em> 07020</th>
</tr>
</thead>
<tbody>
<tr>
<td>g1826.t1</td>
<td>g9539.t1</td>
<td>g4451.t1</td>
<td>g11430.t1</td>
<td>g9867.t1</td>
<td>g4309.t1</td>
<td></td>
</tr>
<tr>
<td>g6500.t1</td>
<td>g3237.t1</td>
<td>g1418.t1</td>
<td>g2957.t1</td>
<td>g11414.t1</td>
<td>g3190.t1</td>
<td></td>
</tr>
<tr>
<td>g10980.t1</td>
<td>g8406.t1</td>
<td>g428.t1</td>
<td>g4027.t1</td>
<td>g1299.t1</td>
<td>g9001.t1</td>
<td></td>
</tr>
<tr>
<td>g6974.t1</td>
<td>g1617.t1</td>
<td>g2727.t1</td>
<td>g6734.t1</td>
<td>g1084.t1</td>
<td>g7673.t1</td>
<td></td>
</tr>
<tr>
<td>g4340.t1</td>
<td>g8457.t1</td>
<td>g6286.t1</td>
<td>g3740.t1</td>
<td>g3117.t1</td>
<td>g5294.t1</td>
<td></td>
</tr>
<tr>
<td>g6884.t1</td>
<td>g9941.t1</td>
<td>g5645.t1</td>
<td>g7583.t1</td>
<td>g6019.t1</td>
<td>g3113.t1</td>
<td></td>
</tr>
<tr>
<td>g4702.t1</td>
<td>g8609.t1</td>
<td>g6364.t1</td>
<td>g2240.t1</td>
<td>g8787.t1</td>
<td>g4479.t1</td>
<td></td>
</tr>
<tr>
<td>g9196.t1</td>
<td>g8787.t1</td>
<td>g1372.t1</td>
<td>g9519.t1</td>
<td>g730.t1</td>
<td>g3953.t1</td>
<td></td>
</tr>
<tr>
<td>g10976.t1</td>
<td>g8461.t1</td>
<td>g10074.t1</td>
<td>g809.t1</td>
<td>g2168.t1</td>
<td>g6740.t1</td>
<td></td>
</tr>
<tr>
<td>g7596.t1</td>
<td>g2716.t1</td>
<td>g3741.t1</td>
<td>g5819.t1</td>
<td>g10432.t1</td>
<td>g2416.t1</td>
<td></td>
</tr>
</tbody>
</table>
Table 3.11 Transposable element sequences downloaded from GenBank that were used in comparisons against the genomes of Microdochium spp.

<table>
<thead>
<tr>
<th>Class</th>
<th>Superfamily*</th>
<th>Family*</th>
<th>Transposon Name</th>
<th>Organism</th>
<th>Accession Number</th>
<th>Function</th>
<th>Sequence Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>LTR retrotransposons</td>
<td>Ty3/gypsy</td>
<td>skippy</td>
<td><em>Fusarium oxysporum</em></td>
<td>AAA88790.1</td>
<td>gag</td>
<td>aa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S60179</td>
<td>pol</td>
<td>aa</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MAGGY</td>
<td><em>Magnaporthe grisea</em></td>
<td>AAA33420.1</td>
<td>pol</td>
<td>aa</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>pyret</td>
<td><em>M. grisea</em></td>
<td>AB062507.1</td>
<td>gag and pol</td>
<td>aa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cgret</td>
<td><em>Colletotrichum. gloeosporoides</em></td>
<td>AAG24792.1</td>
<td>pol</td>
<td>aa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AAG24791.1</td>
<td>gag</td>
<td>aa</td>
</tr>
<tr>
<td>II</td>
<td>TC1 / mariner</td>
<td>Pogo</td>
<td>Fot1</td>
<td><em>F. oxysporum</em></td>
<td>EMT73539.1</td>
<td>unknown</td>
<td>aa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>flipper</td>
<td><em>Botrytis cinerea</em></td>
<td>AAB63315.1</td>
<td>transposase</td>
<td>aa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pat</td>
<td><em>P. anserina</em></td>
<td>unknown</td>
<td>aa</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>crawler</td>
<td><em>Aspergillus oryzae</em></td>
<td>BAE93244.1</td>
<td>transposase</td>
<td>aa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ant1 and Tc1</td>
<td>impala</td>
<td>AAB33090.2</td>
<td>transposase</td>
<td>aa</td>
</tr>
<tr>
<td></td>
<td>hAT-like</td>
<td>restless</td>
<td>Tolypocladium inflatum</td>
<td>CAA93759.1</td>
<td>transposase</td>
<td>aa</td>
<td></td>
</tr>
</tbody>
</table>

*§* indicates non-LTR elements.
<table>
<thead>
<tr>
<th>Class</th>
<th>Superfamily*</th>
<th>Family*</th>
<th>Transposon Name</th>
<th>Organism</th>
<th>Accession Number</th>
<th>Function</th>
<th>Sequence Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mutator-like</td>
<td>-</td>
<td>mutator</td>
<td>Neofusicoccum parvum</td>
<td>EOD48650.1</td>
<td>mutator</td>
<td>aa</td>
<td></td>
</tr>
<tr>
<td>MITE-like</td>
<td>-</td>
<td>Guest</td>
<td>N. crassa</td>
<td>AY374119.1</td>
<td>unknown</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>mimp</td>
<td>F. oxysporum</td>
<td>EU833101.1</td>
<td>mimp4</td>
<td>nt</td>
<td></td>
</tr>
<tr>
<td>Activator</td>
<td>-</td>
<td>VdHAT</td>
<td>Verticillium dahliae</td>
<td>JN160811.1</td>
<td>unknown</td>
<td>nt</td>
<td></td>
</tr>
</tbody>
</table>

* From (Daboussi and Capy 2003)
§ - no family assigned
† - RT = reverse transcriptase
Table 3.12 Summary of putative transposable element sequences identified in the *Microdochium* spp. genomes and their relative proximity to putative pathogen-host interaction genes

<table>
<thead>
<tr>
<th>Species</th>
<th>Isolate</th>
<th>Total number of putative TEs*</th>
<th>Number of TEs found within 5kb of a PHI gene</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>M. nivale</em></td>
<td>11037</td>
<td>328</td>
<td>198</td>
</tr>
<tr>
<td></td>
<td>12262</td>
<td>230</td>
<td>153</td>
</tr>
<tr>
<td></td>
<td>10106</td>
<td>196</td>
<td>121</td>
</tr>
<tr>
<td><em>M. majus</em></td>
<td>99049</td>
<td>240</td>
<td>141</td>
</tr>
<tr>
<td></td>
<td>10095</td>
<td>276</td>
<td>215</td>
</tr>
<tr>
<td><em>M. bolleyi</em></td>
<td>07020</td>
<td>317</td>
<td>162</td>
</tr>
</tbody>
</table>

* putative TEs were identified as unique if they were not found within 500 bp of another match to prevent counting the same match twice.
Figure 3.1 Pipeline for DNA sequencing by Illumina-Solexa technology. Genomic DNA (1) is isolated and (2) sheared into millions of fragments. Adapter sequences are ligated onto both sides of all fragments (3). The fragments are introduced to the solid surface (4), which contains sequences that are complementary to the adapters. In the PCR stage (5), the second adapter sequence of each fragment (white) anneals to its complement. The fragment is amplified by DNA polymerase to produce a complementary strand (6). In each round of PCR, both strands are duplicated (7). The result is clusters of identical complementary sequences (8) for each of the original fragments. In the sequencing stage, a solution containing all four dNTPs, each labelled with a different terminating fluorophore, is introduced (9). Only one nucleotide at a time can be introduced (10) due to the terminator sequence. The incorporated nucleotide is identified by the fluorophore. The fluorophore terminator is cleaved (11) to allow for the addition of the next nucleotide. The labelled dNTP solution is re-applied (12), the second nucleotide is incorporated (13), and step 11 is repeated. This process is repeated until the full fragment has been sequenced (14) (Mardis 2008).
Figure 3.2 Summary of paired-end sequencing. Genomic DNA is sheared into fragments of approximately the same length (e.g. approximately 500 bp). This fragment is then sequenced from both the 5' and the 3' ends, generating two "paired" fragments (labelled A and B and joined by a dashed arc). The number of sequencing cycles performed is equal to the lengths of these fragments (e.g. for strands consisting of two paired-end fragments of 100 bp each, a total of 200 cycles would be performed). During the assembly process, the expected distance between these two fragments (in this example, 300 bp) is used to facilitate their association with the other fragments generated during sequencing.
Figure 3.3 Schematic representation of a de Bruijn graph (Zerbino and Birney 2008). Nodes are represented by boxes. Each node consists of a short alignment of overlapping sequences of the same length. Each node also has a sister node consisting of the reverse complement of the sequences and the alignment in its sister, located immediately above or below the node (e.g. A and A' are sister nodes). Nodes are connected based on their sequence similarities: for example, the final sequence in node A shares four of its six nucleotides with the first sequence in node B. Node B is connected both to node A and to nodes C and D, because the final sequence in node B overlaps equally well with the first sequences in both nodes C and D. The use of de Bruijn graphs results in the association of short, overlapping alignments that are used to assemble the sequencing reads into larger contigs or scaffolds.
Figure 3.4 a) Neighbour-joining, b) maximum likelihood, and c) maximum parsimony trees depicting the relationships between the sequenced *Microdochium* genomes, based on the concatenated sequences of ten genes that were putatively unique to *Microdochium*. Bootstrap values (out of 100) are displayed on each node. The ten genes are listed in Table 3.10. Scale bars represent either 0.1 nucleotide change per base (a and b) or 100 substitutions (c).
Appendices for Chapter 3

Appendix 3.1 Sample script to execute SOAPdenovo

#!/bin/bash
# ./0run_soap.sh 1>soap_25-63.out 2>soap_25-63.errors &

PATH=$PATH:/home/thsiang/programs/soap

for ((i=25; i<64; i+=2))
do
mkdir kmer$i
soap_105_63k all -p 2 -s config.txt -K $i -o ./kmer$i/mm10095_kmer$i
done

for ((i=25; i<64; i+=2))
do
GapCloser -b config.txt -a ./kmer$i/mm10095_kmer$i.scafSeq -o ./kmer$i/mm10095_gclose$i.nt
done
Appendix 3.2 Sample configuration file for SOAPdenovo

```plaintext
# maximal read length
max_rd_len=100

[LIB]
# average insert size
avg_ins=400
# if sequence needs to be reversed
reverse_seq=0
# in which part(s) the reads are used
asm_flags=3
# use only first 100 bps of each read
rd_len_cutoff=100
# in which order the reads are used while scaffolding
rank=1
# cutoff of pair number for a reliable connection (at least 3 for short insert size)
pair_num_cutoff=3
# minimum aligned length to contigs for a reliable read location (at least 32 for short insert size)
map_len=32

# a pair of fastq file, read 1 file should always be followed by read 2 file
q1=/home/thsiang/dna_data/1304_genomes/microdochium_majus10095_R1.fastq.gz
q2=/home/thsiang/dna_data/1304_genomes/microdochium_majus10095_R2.fastq.gz
# a pair of fasta file, read 1 file should always be followed by read 2 file
# f1=/path/**LIBNAMEA**/fasta1_read_1.fa
# f2=/path/**LIBNAMEA**/fasta1_read_2.fa
# fastq file for single reads
# q=/path/**LIBNAMEA**/fastq1_read_single.fq
# fasta file for single reads
# f=/path/**LIBNAMEA**/fastal1_read_single.fa
# a single fasta file for paired reads
# p=/path/**LIBNAMEA**/pairs1_in_one_file.fa
# bam file for single or paired reads, reads 1 in paired reads file should always be followed by reads 2
# NOTE: If a read in bam file fails platform/vendor quality checks (the flag field 0x0200 is set), itself and it's paired read would be ignored.
# b=/path/**LIBNAMEA**/reads1_in_file.bam
```
Appendix 3.3 Sample script used to execute ABySS

#!/bin/bash
# ./abyss.sh -p 2 2>abyss25-63_new.err 1>abyss25-63_new.out &
PATH=$PATH:/home/ljewell/programs/abyss-1.3.5/bin
for ((i=25; i<64; i+=2))
  do
    mkdir kmer$i
    abyss-pe k=$i n=5 name=mb07020_kmer$i in='/home/thsiang/dna_data/1304_genomes/microdochium_bolleyi07020_R1.fastq.gz
    /home/thsiang/dna_data/1304_genomes/microdochium_bolleyi07020_R2.fastq.gz'
done
Appendix 3.4 Sample script used to execute Velvet

```bash
# sqsub --mpp=16g -o 99049_Velvet57-65.out -r 2d bash run_Velvet.sh

PATH=$PATH:/work/ljewell/hound/programs/Velvet/

for ((i=57; i<65; i+=2))
do
    mkdir k$i
    Velveth ./k$i $i -shortPaired ../majus_paired.fas
    Velvetg ./k$i -ins_length 200 -exp_cov 20
done
```
Appendix 3.5 Sample script used to execute AUGUSTUS

#!/bin/bash
#./0run_agustus.sh 1>aug.out 2>aug.err &

/home/ljewell/programs/augustus/bin/augustus --
AUGUSTUS_CONFIG_PATH=/home/ljewell/programs/augustus/config/ --
species=magnaporthe_grisea ../abyss/kmer63/mm10095_kmer63-scaffolds.fa >
scaffPred_mm10095_a63

Perl /home/ljewell/programs/augustus/scripts/getAnnoFasta.pl
scaffPred_mm10095_a63 --seqfile=../abyss/kmer63/mm10095_kmer63-scaffolds.fa
Appendix 3.6 Annotate_genes.pl

#!/usr/bin/perl
#Script to parse through FastAnnotate results to associate them with full
#names from the Pfam table

#script use: ./annotate_genes.pl fa_output pfam_table pred_genes genome
#Where:
#fa_output is the output of fastannotate
#pfam_table is the downloaded table of all pfam accessions
#pred_genes is the file that was used for the fastannotate submission
#and "genome" is the desired genome name (eg isolate number) to be appended
to the name of the gene

open $IN, "ARGV[0]";
open $PFAM, "ARGV[1]";
open $GENOME, "ARGV[2]";
open $OUT, "ARGV[2].annotated";
$genome_name = "ARGV[3]";
%genome = ();
%pfam   = ();

while (<$GENOME>) {
    chomp $_;
    if ($_ =~ />/) {
        $name = $_;
        $name =~ s/>//g;
        $seq = <$GENOME>;
        chomp $seq;
        $genome{$name} = $seq;
    }
}

while (<$PFAM>) {  
    chomp $_;
    @line = split(/\t/, $_);
    $acc = $line[4];
    $acc =~ /PF(\d+)/;
    $acc = $1;
    $pfam{$acc} = $line[6];
}

while (<$IN>) {
    chomp $_;
    @line = split(/\t/, $_);
    $gene = $line[0];
    $pfam_acc = $line[10];

    if ($pfam_acc =~ /-/) {
        print $OUT "$genome_name|$gene| Unknown function
$genome($gene)"
    }
    else {
        $pfam_acc =~ /pfam(\d+)/;
    }
}
$pfam_acc = $1;

if (exists($pfam{$pfam_acc})) {
    $desc = $pfam{$pfam_acc};
    $desc =~ s/\\n//g;
    print $OUT ">$genome_name|$gene| $desc
genome($gene)
"
;
} else {
    print $OUT ">$genome_name|$gene| Unknown function
genome($gene)
"
;
}

close $IN;
close $PFAM;
close $OUT;
close $GENOME;
Appendix 3.7 parse_m9.pl

#!/usr/bin/perl
use warnings;
#parse_m9.pl

#Script to parse formatting from BLAST searches conducted with -m9 flag
#Note that this script accepts a LIST OF FILES as its input
#NOT BLAST outfiles directly
#Script use: ./parse_m9.pl list_of_BLAST_outfiles
#Output is BLAST_outfile.parsed
#(where BLAST_outfile is the name of your raw BLAST output)

open $LIST, "<$ARGV[0]";
$list = <$LIST>;
for ($i=0; $i<=$#list; $i++) {
    chomp $list[$i];
    open $IN, "<$list[$i]";
    open $OUT, "+>$list[$i].parsed";

    while (<$IN>) {
        if ($_ =~ /#/){
        } else {
            print $OUT "$_";
        }
    }
    close $OUT;
    close $IN;
}
close $LIST;
Appendix 3.8 make_simple_table_v2.pl

#!/usr/bin/perl
#13.01.01
use warnings;
# make_simple_table_v2.pl

#Script use:
# ./make_simple_table_v2.pl list names max_eval

#NOTE: max_eval is the MAX exponent allowed, e.g. for e-05 :
#    ./make_simple_table.pl list names 05
#"list" is a file that contains the filenames of the parsed BLAST results
#to make this list: ls *.parsed > list  [for example]
#"names" is a file that contains the NAMES ONLY of all predicted genes

# This script will parse multiple BLAST output files and will create a master
table in the following format:
#
# species 1    species 2    .    .    .
# gene name    hit_exists  hit_exists.....
#
# gene 1    0    1
# gene 2    1    0
# .    .    .

###NOTE###
# This script should be run IN THE FOLDER THAT CONTAINS THE BLAST OUTPUT
# FILES!!
# Blast should be run using the -m 9 flag (outputs results in tabular form)
and parsed to remove the comments (script parse_m9.pl)

#Open the list file:
$listname = $ARGV[0];
chomp $listname;

#Throw the filenames into an array:
open $LIST, "<$listname";
@list = <$LIST>;
close $LIST;

#grab the max evalue from the input line:
$max_eval = $ARGV[2];

#open the list that contains the names of the genes:
open $NAMES, "$ARGV[1]";
@gene_names = <$NAMES>;
close $NAMES;
chomp @gene_names;

#open an outfile and tag it with the evalue for later reference:
open $OUT, "simple_table.$max_eval.txt";

# Print the table headers into the outfile:
print $OUT "Gene_name\t";

# Parse and then print the names of the species that the BLAST was performed against:
foreach $species (@list) {
    chomp $species;
    $temp_name = $species;
    $temp_name =~ s/\..parsed/\n/g;
    print $OUT "$temp_name\t";
}
print $OUT "\n\t";

# Print some human-friendly headers:
$num_list = 0;
until ($num_list > $#list) {
    print $OUT "hit_exists\t";
    $num_list ++;
}
print $OUT "\n";

# Create a hash that will ultimately become the hash-of-hashes that stores all of the data:
%hash = ();

# For each species in turn, open the corresponding BLAST output file:
foreach $sp_name(@list) {
    chomp $sp_name;
    open ($BLAST, "$sp_name");

    # Go line by line through the BLAST output:
    foreach $entry (<$BLAST>) {
        chomp $entry;

        # Split each line. Grab the name (the 0th entry) and the evalue (the 10th entry):
        @line = split(/\s+/, $entry);
        $genename = $line[0];
        $eval = $line[10];

        # If the evalue is EITHER 0.0 OR if the exponent is GREATER than the max evalue (exp >= evalue)
        # Store the name of the gene in the hash for the species, which in turn is stored in the global hash:
        # If this requirement is not met, do nothing.
        # Note that this means that you only need ONE hit that meets the threshold value to get a "positive" result in the final table
        if ($eval =~ /0.0/) {
            $hash{$sp_name}{$genename} = $eval;
        } elsif ($eval =~ /e/) {

        }
}
$eval =~ s/\d+e-//g;
if ($eval >= $max_eval) {
    $hash{$sp_name}{$genename} = $eval;
}

} close $BLAST;

#When all of the data has been stored, parse through it and output the results into a table for further analysis:

#For all of the genes:
foreach $name (@gene_names) {
    #Create a row in the table that starts with the name of the gene:
    print $OUT "$name\t";
    #Now go through each species in the hash of hashes and look for an entry corresponding to the gene of interest:
    foreach $species (@list) {
        #If the entry exists, print a value of 1
        if ($hash{$species}{$name}) {
            print $OUT "1\t";
        }
        #If the entry does not exist, print a value of 0
        else {
            print $OUT "0\t";
        }
    }
    print $OUT "\n";
}
Appendix 3.9 summarize_with_files_v2.pl

#!/usr/bin/perl
#13.05.06
use warnings;

#This is an updated version of my old "summarize_with_files" script
#It takes your binary hit matrix AS WELL AS the species and a host

# Script use: ./summarize_with_files_v2.pl table.txt species host
#e.g. ./summarize_with_files_v2.pl simple_table.txt nivale wheat
#NOTE: host is optional for majus

#Note that this script runs, at its core, as a series of subroutines
#These subroutines are explained more thoroughly within their guts:

$my_species = $ARGV[1];
if ($my_species =~ /nivale/) {
    $host = $ARGV[2];
    if ($host) {
    } else {
        die "ERROR: Please enter a host name for M. nivale (wheat or grass)\n";
    }
} elsif ($my_species =~ /majus/) {
    $host = "wheat";
}
&determine_relationships();
&open_files();
&define_variables();
$gene_counter = 0;
while (<$IN>) {
    chomp $;
    # Check the current line. If it is the first line, it contains a list of
    # the names which much be parsed to determine which data is where
    # Split the line into an array and examine the second column. If it is
    # the first line, the second column will contain non-digit characters (i.e.
    # If it is a data line, it will contain either a 0 or 1 (\d)

    @temp_array = split (/\s+/, $);
    if ($temp_array[1] =~ /hit/) {
    }
    elsif ($temp_array[1] =~ /\D/) {
        &identify_columns();
    }
}
else {
    $gene_counter++;
    &determine_relationships();
    &do_math();
    &parse_results();
}

&print_outfiles();
&close_files();

#############
#SUBROUTINES#
#############

sub do_math {
    $num_genes ++;
    $total = 0;
    $gene_name = $temp_array[0];

    #Tally the results (i.e. the 1's and 0's) to determine the total number
    #of times that a particular gene was found:
    for ($z = 1; $z <= $#temp_array; $z ++) {
        $total += $temp_array[$z];
    }

    #Immediately check if the sum is zero. If it is, there is no need to
    #perform any of the other checks:
    if ($total == 0) {
        $never++;
        print $NEVER "$gene_name
    next;
}

    #Determine the "exclusivity" of the hit:
    $sister_sum = 0;
    foreach $sis (@sister_locations) {
        $sister_sum+=$temp_array[$sis];
    }

    $intra_sum = 0;
    foreach $intra (@intra_locations) {
        $intra_sum += $temp_array[$intra];
    }

    $wheat_sum = 0;
    foreach $wheat (@wheat_locations) {
        $wheat_sum += $temp_array[$wheat];
    }

    $grass_sum = 0;
    foreach $grass (@grass_locations) {
        $grass_sum += $temp_array[$grass];
    }
}
$mb_sum = 0;
foreach $mb (@mb_locations) {
    $mb_sum += $temp_array[$mb];
}

$sall_mic_sum = $sister_sum + $intra_sum + $mb_sum;

$other_sum = 0;
foreach $sord (@other_locations) {
    $other_sum += $temp_array[$sord];
}

$maj_sum = 0;
foreach $maj (@majus_locations) {
    $maj_sum += $temp_array[$maj];
}

$niv_sum = 0;
foreach $niv (@nivale_locations) {
    $niv_sum += $temp_array[$niv];
}

$xyl_sum = 0;
foreach $xyl (@xyl_locations) {
    $xyl_sum += $temp_array[$xyl];
}

$mm99049_sum = 0;
foreach $mm99049 (@mm99049_locations) {
    $mm99049_sum += $temp_array[$mm99049];
}

$mm10095_sum = 0;
foreach $mm10095 (@mm10095_locations) {
    $mm10095_sum += $temp_array[$mm10095];
}

$mn12262_sum = 0;
foreach $mn12262 (@mn12262_locations) {
    $mn12262_sum += $temp_array[$mn12262];
}

$mn10106_sum = 0;
foreach $mn10106 (@mn10106_locations) {
    $mn10106_sum += $temp_array[$mn10106];
}

$mn11037_sum = 0;
foreach $mn11037 (@mn11037_locations) {
    $mn11037_sum += $temp_array[$mn11037];
}

sub parse_results {
    #First some general stuff:
    if ($mm99049_sum != 0) {
        $total_in_99049++;
if ($mm10095_sum != 0) {
    $total_in_10095++;
}
if ($mn12262_sum != 0) {
    $total_in_12262++;
}
if ($mn11037_sum != 0) {
    $total_in_11037++;
}
if ($mn10106_sum != 0) {
    $total_in_10106++;
}

if (($my_species =~ /niva/) || ($my_species =~ /majus/)){
    if ($intra_sum != 0) {
        $total_in_intra++;
    }
    if ($sister_sum != 0) {
        $total_in_sister++;
    }
    if ($mb_sum != 0) {
        $total_in_mb++;
    }
}
elsif (($my_species =~ /bolleyi/)){
    if ($niv_sum != 0) {
        $total_in_niv++;
    }
    if ($maj_sum != 0) {
        $total_in_maj++;
    }
}
if ($other_sum != 0) {
    $total_in_sord++;
}
if ($xyl_sum != 0) {
    $total_in_xyl++;
}

#if there are no "other" hits (i.e. no hits in a non-xylariales...)
if ($other_sum == 0) {
    #If there are no hits in the NON-MIC xylariales, this means that the
    only hits are among the mic species
    $not_in_sord++;
}

if ($xyl_sum != 0) {

$in_any_xyl++; 
if (($niv_sum == 0) && ($maj_sum == 0) && ($mb_sum == 0)) {
    print $OTHER_XYL_ONLY "$gene_name\n";
    $in_other_xyl_only++;
}
}

if (($my_species =~ /nivale/) || ($my_species =~ /majus/)) {

    if (($xyl_sum == 0) && ($mb_sum == 0)) {
        if (($intra_sum == 0) && ($sister_sum != 0)) {
            print $SISTER "$gene_name\n";
            $in_sister_only++;
        }
        if (($intra_sum != 0) && ($sister_sum == 0)) {
            print $INTRA "$gene_name\n";
            $in_intra_only++;
        }
        if (($sister_sum != 0) && ($intra_sum != 0)) {
            print $NIV_MAJ_ONLY "$gene_name\n";
            $in_niv_and_maj_only++;
        }
    }
    if (($grass_sum != 0) && ($wheat_sum == 0)) {
        print $NIV_GRASS "$gene_name\n";
        $only_niv_grass++;
    }
    if (($grass_sum == 0) && ($wheat_sum != 0)) {
        print $NIV_WHEAT "$gene_name\n";
        $only_niv_wheat++;
    }
    if (($xyl_sum == 0) && ($sister_sum == 0) && ($intra_sum == 0) && ($mb_sum != 0)) {
        print $MB "$gene_name\n";
        $in_mb_only++;
    }
    if (($xyl_sum == 0) && ($sister_sum != 0) && ($intra_sum != 0) && ($mb_sum != 0)) {
        print $MIC_ONLY "$gene_name\n";
        $in_all_mic_only++;
    }
    if (($xyl_sum != 0) && ($sister_sum != 0) && ($intra_sum != 0) && ($mb_sum != 0)) {
        print $ALL_XYL "$gene_name\n";
        $in_all_xyl++;
    }
}

181
elsif ($my_species =~ /bolleyi/)
    if (($xyl_sum == 0) && ($niv_sum != 0) && ($maj_sum != 0)) {
        print $SISTER "\$gene_name\n";
        $in_sister_only++;
    }
    if (($xyl_sum == 0) && ($maj_sum != 0) && ($niv_sum == 0)) {
        $in_maj_only++;
    }
    if (($xyl_sum == 0) && ($maj_sum == 0) && ($niv_sum != 0)) {
        $in_niv_only++;
    }
    if (($xyl_sum != 0) && ($niv_sum != 0) && ($maj_sum != 0)) {
        print $ALL_XYL "\$gene_name\n";
        $in_all_xyl++;
    }
}
else {
    if (($my_species =~ /nivale/) || ($my_species =~ /majus/)) {
        if (($xyl_sum != 0) && ($intra_sum != 0) && ($sister_sum != 0) && ($mb_sum != 0)) {
            print $IN_ALL "$\$gene_name\n";
            $in_all++;
        }
        if (($xyl_sum != 0) && ($intra_sum != 0) && ($mb_sum != 0) && ($sister_sum == 0)) {
            print $NOT_IN_SIS "$\$gene_name\n";
            $not_in_sister++;
        }
        if (($xyl_sum != 0) && ($intra_sum == 0) && ($mb_sum != 0) && ($sister_sum != 0)) {
            print $NOT_IN_INTRA "$\$gene_name\n";
            $not_in_intra++;
        }
    }
    elsif ($my_species =~ /bolleyi/) {
        if (($xyl_sum != 0) && ($niv_sum != 0) && ($maj_sum != 0)) {
            print $IN_ALL "$\$gene_name\n";
            $in_all++;
        }
    }
}
if (($xyl_sum == 0) && ($all_mic_sum == 0)) {
    print $SORD_ONLY "$gene_name
    $in_sord_only++;
}
if (($xyl_sum != 0) && ($all_mic_sum == 0)) {
    print $IN_OTHER_ONLY "$gene_name
    $in_other_only++;
}
}

sub determine_relationships {
    #The locations of the species of interest were determined by the
    "identify_columns" subroutine

    $my_species = $ARGV[1];
    if ($my_species =~ /majus/) {
        $sister_name = "nivale";
        @intra_locations = @majus_locations;
        @sister_locations = @nivale_locations;
    }
    elsif ($my_species =~ /nivale/) {
        $sister_name = "majus";
        @intra_locations = @nivale_locations;
        @sister_locations = @majus_locations;
    }
    elsif ($my_species =~ /bolleyi/) {
        $sister_name = "niv_and_maj";
        @intra_locations = ();
        @sister_locations = ();
    }
    else {
        die "ERROR: species name was not found. Please enter the species name
        (majus, nivale, or bolleyi) in lowercase letters.\n"
    }
}

sub identify_columns {
    &define_groups();

    @majus_locations = ();
    @nivale_locations = ();
    @mb_locations = ();
    @xyl_locations = ();
    @other_locations = ();
@wheat_locations = ();
@grass_locations = ();

# Start at 1 (not 0) because first column is "gene_name"
for ($a = 0; $a <= $#temp_array; $a++) {
    foreach $sp (@xylariales) {
        if ($temp_array[$a] =~ /$sp/) {
            push (@xyl_locations, $a);
        }
    }
    foreach $sp (@mb) {
        if ($temp_array[$a] =~ /$sp/) {
            push (@mb_locations, $a);
        }
    }
    foreach $sp (@majus) {
        if ($temp_array[$a] =~ /$sp/) {
            push (@majus_locations, $a);
        }
    }
    foreach $sp (@nivale) {
        if ($temp_array[$a] =~ /$sp/) {
            push (@nivale_locations, $a);
        }
    }
    foreach $sp (@wheat) {
        if ($temp_array[$a] =~ /$sp/) {
            push (@wheat_locations, $a);
        }
    }
    foreach $sp (@grass) {
        if ($temp_array[$a] =~ /$sp/) {
            push (@grass_locations, $a);
        }
    }
    foreach $sp (@other) {
        if ($temp_array[$a] =~ /$sp/) {
            push (@other_locations, $a);
        }
    }
    foreach $sp (@mm99049) {
        if ($temp_array[$a] =~ /$sp/) {
            push (@mm99049_locations, $a);
        }
    }
    foreach $sp (@mm10095) {
        if ($temp_array[$a] =~ /$sp/) {
            push (@mm10095_locations, $a);
        }
    }
    foreach $sp (@mn11037) {
if ($temp_array[$a] =~ /$sp/) {
    push (@mn1037_locations, $a);
}

foreach $sp (@mn12262) {
    if ($temp_array[$a] =~ /$sp/) {
        push (@mn12262_locations, $a);
    }
}

foreach $sp (@mn10106) {
    if ($temp_array[$a] =~ /$sp/) {
        push (@mn10106_locations, $a);
    }
}

sub print_outfiles {
    # This prints a quick, human-readable summary of the results:
    print $OUT "For $ARGV[0]: \n";
    print $OUT "########

    print $OUT "A total of $gene_counter genes were assessed.\n"
    print $OUT "A total of $never genes had no match in any of the genomes studied.\n"
    print $OUT "There were $in_all genes that were in ALL categories.\n"
    print $OUT "There were $not_in_sord genes NOT in the non-xylariales sordariales.\n"
    print $OUT "There were $in_any_xyl genes NOT in a sordariales that WERE in the xylariales (either unique to xyl, unique to mic, or not shared at all).\n"
    print $OUT "There were $in_other_xyl_only genes that were ONLY found among the NON-Microdochium Xylariales.\n"
    print $OUT "There were $in_all_xyl genes that were in ALL of the xylariales, including the Microdochium species\n"
    print $OUT "There were $in_other_only genes that were only found in the NON-Microdochium species.\n"
    print $OUT "There were $in_all_xyl only genes that were ONLY found among the NON-Microdochium Xylariales.\n"
    print $OUT "There were $in_niv_and_maj_only genes that were only found in BOTH M. nivale and M. majus but no other species.\n"
    print $OUT "There were $in_mb_only genes that were only found in M. bolleyi.\n"
    print $OUT "There were $in_sister_only genes in the sister species ($sister_name) ONLY.\n"
    print $OUT "There were $in_intra_only genes that were found only in intraspecific genomes ($my_species) ONLY.\n"
Among ALL of the hits shared with M.nivale:


There were $only_niv_wheat genes that were shared ONLY with the wheat isolates.


There were $only_niv_grass genes that were shared ONLY with the grass isolates.


There were $not_in_intra genes that were found in every species EXCEPT the intra-specific genomes ($my_species).


There were $not_in_sister genes that were found in every species EXCEPT the sister species genomes ($sister_name).


There were $in_niv_only genes found ONLY among the M. nivale isolates.


There were $in_maj_only genes found ONLY among the M. majus isolates.


There were $total_in_niv TOTAL genes found among the M. nivale isolates.


There were $total_in_maj TOTAL genes found among the M. majus isolates.


There were $total_in_sister TOTAL genes found among the $sister_name isolates.


There were $total_in_intra TOTAL genes found among the $my_species isolates.


There were $total_in_mb TOTAL genes found in M. bolleyi.


There were $total_in_10095 TOTAL genes found in M. majus 10095.


There were $total_in_11037 TOTAL genes found in M. nivale 11037.


There were $total_in_12262 TOTAL genes found in M. nivale 12262.


There were $total_in_10106 TOTAL genes found in M. nivale 10106.


There were $total_in_99049 TOTAL genes found in M. majus 99049.


There were $total_in_sord TOTAL genes found among the non-Xylariales and non-Microdochium Sordariomycetes.


There were $total_in_xyl TOTAL genes found among the non-Microdochium Xylariales.

}
### Basic file handle stuff ###

```perl
sub open_files {
  open ($IN, "$ARGV[0]");
  open ($OUT, "+>$ARGV[0].summary");

  open ($MIC_ONLY, "+only_in_microdochium.txt");
  open ($INTRA, "+only_in_$my_species.txt");
  open ($SISTER, "+only_in_$sister_name.txt");
  open ($NIV_MAJ_ONLY, "+only_in_nIv_and_maj.txt");
  open ($MB, "+only_in_mb.txt");
  open ($NIV_WHEAT, "+niv_wheat_only.txt");
  open ($NIV_GRASS, "+niv_grass_only.txt");

  open ($NEVER, "+never_found.txt");
  open ($IN_ALL, "+in_all_groups.txt");
  open ($SORD_ONLY, "+not_in_any_xyl.txt");
  open ($OTHER_XYL_ONLY, "+only_in_other_xyl.txt");
  open ($ALL_XYL, "+in_all_xyl.txt");
  open ($IN_OTHER_ONLY, "+in_non_mic_only.txt");

  open ($NOT_IN_SIS, "+not_in_$sister_name.txt");
  open ($NOT_IN_INTRA, "+not_in_$my_species.txt");
}

sub close_files {
  close $IN;
  close $OUT;

  close $MIC_ONLY;
  close $INTRA;
  close $SISTER;
  close $NIV_MAJ_ONLY;
  close $MB;
  close $NIV_WHEAT;
  close $NIV_GRASS;

  close $NEVER;
  close $IN_ALL;
  close $SORD_ONLY;
  close $OTHER_XYL_ONLY;
  close $ALL_XYL;
  close $IN_OTHER_ONLY;

  close $NOT_IN_SIS;
  close $NOT_IN_INTRA;
}
```

##### Define groups #####

---

187
sub define_groups {
    @xylariales = (
        "astygium.fa",
        "deschscholzii.fa",
        "hypoxylon_co.fa",
        "ptheae.nt",
    );

    @nivale = (
        "mn10106_s63.80",
        "mn12262_abyss53",
        "mn11037_abyss59",
    );

    @wheat = (
        "mn10106_s63.80",
        "mn12262_abyss53",
    );

    @grass = (
        "mn11037_abyss59",
    );

    @majus = (
        "mm10095_a63.nt",
        "mm99049_abyss45",
    );

    @mb = (
        "mb07020_s61.80.nt",
    );

    @other = (
        "aalcalophilum.nt",
        "cgram.nt",
        "chaetom.nt",
        "fusoxyl.nt",
        "fusvert.nt",
        "gclavigera.nt",
        "moryzae.nt",
        "mpoae.nt",
        "ncrassa.nt",
        "sthermophile.nt",
        "ttterrestris.nt",
        "valboatrum.nt",
        "vdahliae.nt",
        "vlongispor.nt",
    );

    @mn12262 = (
        "mn12262_abyss53",
    );

    @mn10106 = (
        "mn10106_s63.80",
    );
}
@mn11037 = ( "mn11037_abyss59",
);
@mm10095 = ( "mm10095_a63.nt",
);
@mm99049 = ( "mm99049_abyss45",
);

###########
sub define_variables {
    $num_genes = 0;
    $in_sister_only = 0;
    $in_all = 0;
    $in_other_only = 0;
    $never = 0;
    $in_mb_only = 0;
    $in_niv_and_maj_only = 0;
    $in_intra_only = 0;
    $not_in_sister = 0;
    $not_in_intra = 0;
    $not_in_sord = 0;
    $in_other_xyl_only = 0;
    $in_any_xyl = 0;
    $in_all_xyl = 0;
    $in_all_mic_only = 0;
    $in_sord_only = 0;
    $in_other_only = 0;
    $only_niv_wheat = 0;
    $only_niv_grass = 0;
    $in_niv_only = 0;
    $in_maj_only = 0;
    $total_in_mb = 0;
    $total_in_sister = 0;
    $total_in_intra = 0;
    $total_in_niv = 0;
    $total_in_maj = 0;
    $total_in_xyl = 0;
    $total_in_sord = 0;
    $total_in_99049 = 0;
    $total_in_10095 = 0;
    $total_in_11037 = 0;
    $total_in_12262 = 0;
    $total_in_10106 = 0;
}

Appendix 3.10 find_genes_of_diff_length.pl

#!/usr/bin/perl

#Script to identify BLAST hits that are both a) highly conserved, and b) differ in length
#Script use:
# ./find_genes_of_diff_length.pl BLAST_output.parsed
#NOTE: BLAST should be run with the m-9 flag and results "pre-parsed" with the parse_m9 script

open $IN, "$ARGV[0]";
open $OUT, "+$ARGV[0].diff_length";

#Print a human-friendly header:
print $OUT "Queryname	Hitname	Diff	Qlength	Hlength	Match_length\n";

#Query id, Subject id, % identity, alignment length, mismatches, gap openings, q. start, q. end, s. start, s. end, e-value, bit score
while (<$IN>) {
    #line by line, look through the BLAST output and examine first the evalue:
    chomp $_;
    @line = split(/\t/, $_);
    # Evalue is entry 10; if the evalue is 0.0, do the length calculation:
    if ($line[10] =~ /0.0/) {
        &calculate();
    }
    # Alternatively, if the evalue is <= e-50 (i.e. if the exponent is >= 50), do the calculation:
    elsif ($line[10] =~ /e-d+/) {
        $line[10] =~ /e-(\d+)/;
        if ($1 >= 50) {
            &calculate();
        }
    }
}
close $IN;
close $OUT;

# To determine the length difference, calculate the length of the query and the length of the hit:
sub calculate {
    $qlength = abs($line[7] - $line[6]);
    $hlength = abs($line[9] - $line[8]);
    # If the absolute value of the difference between these lengths is greater than 50, grab this data for further analysis:
    $difference = abs($qlength - $hlength);
    if ($difference > 50) {
        print $OUT "$line[0]\t$line[1]\t$difference\t$qlength\t$hlength\t$line[3]\n";
    }
}
Appendix 3.11 compare_phi_results.pl

#!/usr/bin/perl
use warnings;

#Script to compare the results of the phi BLAST jobs

#script use: ./compare_phi_results.pl list phi.fasta
#where list is a list of all the parsed BLAST files and phi.fasta is all of
#the protein sequences searched

open $LIST, "$ARGV[0]";
@list = <$LIST>;
close $LIST;

&open_infiles();
&open_outfiles();

open $PHI, "$ARGV[1]";
%names = ();
while (<$PHI>) {
    chomp $_
    if ($_ =~ />/) {
        $_ =~ s/>//g;
        $names{$_} = $_;
    }
}
close $PHI;

#doing a hash first will strip duplicates...
@names = keys %names;
&do_comparison();
&close_outfiles();

sub do_comparison {
    foreach $name (@names) {
        if ( ($mn12262{$name}) && ($mn10106{$name}) && ($mn11037{$name}) &&
            ($mm99049{$name}) && ($mm10095{$name}) && ($mb07020{$name})) {
            print $IN_ALL "$name
        }
        elsif ( ($mn12262{$name}) && ($mn10106{$name}) && ($mn11037{$name})
            && ($mm99049{$name}) && ($mm10095{$name}) ) {
            unless ($mb07020{$name}) {
                print $NOT_MB "$name
            }
        }
        elsif ( ($mm99049{$name}) && ($mm10095{$name}) ) {
            
        }
    }
}


unless ( ($mn12262{$name}) || ($mn10106{$name}) || ($mn11037{$name}) || ($mb07020{$name}) ){
    print $MAJ_ONLY "$name\n";
}

elsif ( ($mn12262{$name}) && ($mn10106{$name}) && ($mn11037{$name}) ) {
    unless ( ($mn99049{$name}) || ($mn10095{$name}) || ($mb07020{$name}) ){
        print $NIV_ONLY "$name\n";
    }
}

elseif ( ($mn12262{$name}) ) {
    unless ( ($mn11037{$name}) || ($mn99049{$name}) || ($mn10095{$name}) || ($mb07020{$name}) ) {
        print $MN12262_ONLY "$name\n";
    }
}

elsif ( ($mn10106{$name}) ) {
    unless ( ($mn11037{$name}) || ($mn99049{$name}) || ($mn10095{$name}) || ($mb07020{$name}) ) { 
        print $MN10106_ONLY "$name\n";
    }
}

elsif ( ($mn11037{$name}) ) { 
    unless ( ($mn10106{$name}) || ($mn99049{$name}) || ($mn10095{$name}) || ($mb07020{$name}) ) { 
        print $MN11037_ONLY "$name\n";
    }
}

elsif ( ($mm99049{$name}) ) { 
    unless ( ($mn10106{$name}) || ($mn11037{$name}) || ($mn12262{$name}) || ($mn10095{$name}) || ($mb07020{$name}) ) { 
        print $MM99049_ONLY "$name\n";
    }
}

elsif ( ($mm10095{$name}) ) { 
    unless ( ($mn10106{$name}) || ($mn11037{$name}) || ($mn12262{$name}) || ($mm99049{$name}) || ($mb07020{$name}) ) { 
        print $MM10095_ONLY "$name\n";
    }
}

elsif ( ($mn99049{$name}) ) { 
    unless ( ($mn10106{$name}) || ($mn11037{$name}) || ($mn12262{$name}) || ($mn10095{$name}) || ($mb07020{$name}) ) { 
        print $MN2262_ONLY "$name\n";
    } 
}

elsif ( ($mn10095{$name}) ) { 
    unless ( ($mn10106{$name}) || ($mn11037{$name}) || ($mn12262{$name}) || ($mn99049{$name}) || ($mb07020{$name}) ) { 
        print $MN10095_ONLY "$name\n";
    } 
}
elsif ( ($mb07020{$name}) ) {
    unless ( ($mn10106{$name}) || ($mn11037{$name}) || ($mm99049{$name}) || ($mm10095{$name}) || ($mn12262{$name}) ) {
        print $MB07020_ONLY "$name
    }
} else {
    unless ( ($mn12262{$name}) || ($mn10106{$name}) || ($mn11037{$name}) || ($mm99049{$name}) || ($mm10095{$name}) || ($mb07020{$name}) ) {
        print $NEVER "$name
    }
}

$tot_11037 = keys %mn11037;
$tot_07020 = keys %mb07020;
$tot_12262 = keys %mn12262;
$tot_10095 = keys %mm10095;
$tot_10106 = keys %mn10106;
$tot_99049 = keys %mm99049;

print $TOTALS "There were $tot_11037 in 11037, $tot_07020 in 07020, $tot_12262 in 12262, $tot_10106 in 10106, $tot_10095 in 10095, and $tot_99049 in 99049 \n";

sub close_outfiles {
    close $MN11037_ONLY;
    close $MN12262_ONLY;
    close $MN10106_ONLY;
    close $MM99049_ONLY;
    close $MM10095_ONLY;
    close $MB07020_ONLY;
    close $NEVER;
    close $NOT_MB;
    close $IN_ALL;
    close $MAJ_ONLY;
    close $NIV_W_ONLY;
    close $NIV_ONLY;
    close $TOTALS;
}
sub open_outfiles {
    open $MN11037_ONLY, ">+in_mn11037_only";
    open $MN12262_ONLY, ">+in_mn12262_only";
    open $MN10106_ONLY, ">+in_mn10106_only";
    open $MM99049_ONLY, ">+in_mm99049_only";
    open $MM10095_ONLY, ">+in_mm10095_only";
sub open_infiles {
    foreach $file (@list) {
        open $IN, "$file";
        if ($file =~ /12262/) {
            %mn12262 = ();
            while (<$IN>) {
                @in = split('/\t/, $');
                $mn12262{ $in[0] } = $in[0];
            }
        } elsif ($file =~ /10106/) {
            %mn10106 = ();
            while (<$IN>) {
                @in = split('/\t/, $');
                $mn10106{ $in[0] } = $in[0];
            }
        } elsif ($file =~ /11037/) {
            %mn11037 = ();
            while (<$IN>) {
                @in = split('/\t/, $');
                $mn11037{ $in[0] } = $in[0];
            }
        } elsif ($file =~ /99049/) {
            %mm99049 = ();
            while (<$IN>) {
                @in = split('/\t/, $');
                $mm99049{ $in[0] } = $in[0];
            }
        } elsif ($file =~ /10095/) {
            %mm10095 = ();
            while (<$IN>) {
                @in = split('/\t/, $');
                $mm10095{ $in[0] } = $in[0];
            }
        } elsif ($file =~ /07020/) {
            %mb07020 = ();
            while (<$IN>) {
                @in = split('/\t/, $');
                $mb07020{ $in[0] } = $in[0];
            }
        }
    }
}
} close $IN;
}
Appendix 3.12 check_proximity.pl

#!/usr/bin/perl
#script use: ./check_proximity.pl phi_output te_output 5000
#Where the final argument is the desired maximum distance between the hits to
#be counted as a "positive" result
#note that blast jobs should be performed using the -m9 flag to produce a
#tabular output

open $PHI, "$ARGV[0]";
open $TE, "$ARGV[1]";
$distance = $ARGV[2];
$ARGV[0] =~ s/_vs_phi_m9//g;

open $OUT, "+>$ARGV[0].compared.$distance";
#Print a human-friendly header:
print $OUT "scaff_name|tphi_hit|tphi_start|tphi_end|tte_name|tte_start|tte_end\n";

#Create a hash that will store the information about the locations of the
TEs:
%te = ();

#Read through the output file for the TE blast results.
#Put the results in a hash, where the keys are the scaffolds with matches,
#and the values are arrays
#Each element of the array consists of the TE name and its start and end
#positions in the scaffold.
while ($line = <$TE>) {
    chomp $line;
    #For each non-comment line, split the line at the tabs to create an
array:
    unless ($line =~ /\#/) {
        @line = split(\t, $line);

        #Ensure that the start and end positions are in the desired order
(may be inveted $:
            if ($line[8] > $line[9]) {
                $location = "$line[0]\t$line[9]\t$line[8]\t"
            }
            else {
                $location = "$line[0]\t$line[8]\t$line[9]\t"
            }

            #If at least one match has already been recorded on this scaffold,
add this to the$
            if ( exists($te{$line[1]} ) ) {
                push @{$te{$line[1]}}, $location);
            } else {
                $te{$line[1]} = [$location,];
            }
        }  #End unless
    }
}  #End while

196
#For the pathogen-host interaction gene hits (PHI hits), read through each non-comment line:
while (<$PHI>) {
  chomp $_;
  unless ($_ =~ /^#/) {
    @entry = split('/\t/', $_);
    $hit_name = $entry[1];
    if ($entry[8] > $entry[9]) {
      $start = $entry[9];
      $end = $entry[8];
    } else {
      $start = $entry[8];
      $end = $entry[9];
    }
    #If there is at least one TE match on the same scaffold as the current PHI scaff, then:
    if (exists($te{$hit_name})) {
      foreach (@{$te{$hit_name}}) {
        @temp = split('/\t/', $_);
        #If the start of one match is within 10 kb of the end of the other, print:
        if ( (abs($start - $temp[2]) <= $distance) || (abs($end - $temp[1]) <= $distance) ) {
          print $OUT "$hit_name\t$entry[0]\t$start\t$end\t$temp[0]\t$temp[1]\t$temp[2]\n";
        }
      }
    }
  } else {
    print $PHI;
  }
}
close $PHI;
close $TE;
close $OUT;
Appendix 3.13 eliminate_duplicates.pl

open $IN, "ARGV[0]";
open $OUT, ">$ARGV[0].no_duplicates2";

%hits = ();
while (<$IN>) {
    chomp $_;
    @line = split(/\t/, $_);

    $te = $line[4];
    $te =~ /w+_\w+/;
    $te = $1;

    $base = 10000;
    $rounded = int($line[6]/$base+1) * $base;

    $scaf = $line[0];
    $concat = $scaf . $te . $rounded;
    $hits{$concat} = $_;
}

print $OUT "$hits{$_}\n" for (keys %hits);

close $IN;
close $OUT;
Appendix 3.14 Alignment of putative *M. bolleyi* 07020 EF-1α sequence with the *Microdochium* EF-1α primers EFNivF and EFMajF and the reverse complement of primer EFMicR

| g6339.t1 | ATGGGTAATCCGACACGGCTCACATCAACGCTCGTTATCGGCCCCAGTCTGATTCCGGC |
| g6339.t1 | AAGTCCACCACCACCGTGTCACCTGATCTAAGTGCCTGTTATCGGACAGGCTACCAC |
| g6339.t1 | GAGAAGTTGCAGAAGAACTGCGAGCTCGGCAAGGTGGTCTGTAGATCGTGCTGGTT |
| g6339.t1 | CTTGACAGCTCAAGGCCAGCTGAGGTGATTATCGGACATCGTGGCGTCAAGGAGG |
| g6339.t1 | TTTGAGACCTCCCGTCAACAGCTGCTGCTACCTCGGCTGACAGGACATCGTGGC |
| g6339.t1 | AAGAACATGATCAGTCTGCCATCGCCGATTCGCGATTCATTTGCTGCGCTGATT |
| g6339.t1 | ACTGTTGAGTCGAGGGTGTTATTCATGCTTCAGGAGGAGGACCGGCGACCCGCTG |
| g6339.t1 | GCCTACACCCCTCCTGAAGCTCCACCAAGGCTGAGCCGACACCCCGTTCGATG |
| g6339.t1 | TGGTACGATCTCTGTTCCAGGAGATCAAGGAGCTTCTCCTGCTACAGAAGGTC |
| g6339.t1 | GGCTACACCCCAACGAGTCTGCTGCCATTTCCGCTTCAAGCCGACAACATG |
| g6339.t1 | CTTTGCACCTCCACCAACTGCCCCCTGTAGTCAAGGCTGCTGAGT |
| g6339.t1 | ATGCTCCTCGGCAAGGCTCTCTTCAGGCCCCATCGGAGCCGACCTGCC |
| g6339.t1 | TCCGACAAGCCCTCCTGCCCTTCCCTCCAGGATGTCTACAAGGATCGGATTGTCGAGG |
| g6339.t1 | GTGCCCCGCGGCCTATCGAGACGGCTACCATCAAGGCGCGACATGGCTGCTACCTCGC |
| g6339.t1 | CCCGCTGTTGCTACCACTGAACTGAGTCTCCGCTGAGTCCACAGACCTTCTCCTCAG |
| g6339.t1 | GCTTTCCCAGTGCAACGACGCTCGGCTCCATCGGCTGAGGCTTCCGCTGACGATTCG |
| g6339.t1 | CGTGGCAACTTGGCCTGATGACCAAGAGGCGACCCCTGTTGGCCGCAACACTTACC |
| g6339.t1 | GCCCCAGTCTACGCTTGAGACCACCGGTCGGGGCGGTACCACCGGGGGCTGTCGATA |
| g6339.t1 | GACTGACACTGCCCCACATTGTGCAAGTTCCAGGACTCTCGGAGGAGGAGTCGACC |
| g6339.t1 | GACTGCCACACTGCCACATTTGCAAGTTACCAGGACTCTCGGAGGAGGAGTCGACCCC |
| g6339.t1 | CAGTTCTCAGGTGTTGAG |
| g6339.t1 | CGTACGAGTACGCTCCACGAGCCATGCTGGTTCGAGGCTGTGAGGCTGAGGAGAGGAGTCGAGG |
| g6339.t1 | GCTACCGAGTACGCTCCACGAGCCATGCTGGTTCGAGGCTGTGAGGAGAGGAGTCGAGG |
| 199 |
g6339.t1  GGCGTTTCCCGTCCGTCATGAGACAGACCGTCGCTGTCGGTGTCATCAAGGCCGTC

g6339.t1  GACAAGTCGGAGACTCTGGCGGGGAAGGACCTGCTGCTGAGAAAGGCTGCGAAGAGG
Appendix 3.15 Alignment of predicted gene sequences that are putatively unique to *Microdochium* spp.

<table>
<thead>
<tr>
<th></th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mn12262</td>
<td>ATGTTTCTCC---------ACCTCCGCCCTCGGGCTGTCCTCCGCCGCCCTCGCTGTGTC</td>
</tr>
<tr>
<td>Mn10095</td>
<td>ATGTTTCTCCCTCTCGTACCTCGCCCTCGGGCTGTCACCTCGGGACACTCGGCTGTC</td>
</tr>
<tr>
<td>Mn99049</td>
<td>ATGTTTCTCCCTCCTACCTCGCCCTCGGGTGGTCTCCGCCGCCACTCGACTCGGCTGTC</td>
</tr>
<tr>
<td>Mn11037</td>
<td>ATGTTTCTCCCTGCTCCGCTCCGCTCGGGTGGTCTCCGCCGCCACTCGACTCGGCTGTC</td>
</tr>
<tr>
<td>Mn10106</td>
<td>ATGTTTCTCC---------ACCTCCGCCCTCGGGCTGTCCTCCGCCGCCCTCGCTGTGTC</td>
</tr>
</tbody>
</table>

| Mb07020  | GCCATCCTCTGCAATTCTGGGAGGCCGAAGCCCCCAAGGACTGCGGCGCCATTGTGAAG          |

| Mn12262  | GCCATCCTCTGCAATTCTGGGAGGCCGAAGCCCCCAAGGACTGCGGCGCCATTGTGAAG          |
| Mn10095  | GCCATCCTCTGCAATTCTGGGAGGCCGAAGCCCCCAAGGACTGCGGCGCCATTGTGAAG          |
| Mn99049  | GCCATCCTCTGCAATTCTGGGAGGCCGAAGCCCCCAAGGACTGCGGCGCCATTGTGAAG          |
| Mn11037  | GCCATCCTCTGCAATTCTGGGAGGCCGAAGCCCCCAAGGACTGCGGCGCCATTGTGAAG          |
| Mn10106  | GCCATCCTCTGCAATTCTGGGAGGCCGAAGCCCCCAAGGACTGCGGCGCCATTGTGAAG          |

| Mb07020  | GCCATCCTCTGCAATTCTGGGAGGCCGAAGCCCCCAAGGACTGCGGCGCCATTGTGAAG          |

| Mn12262  | AACATGCAGTACGAGGAACTCAACAACAAGGGACC---GTGGGACATTTTTCTGGGCGGCTG      |
| Mn10095  | AACATGCAGTACGAGGAACTCAACAACAAGGGACC---GTGGGACATTTTTCTGGGCGGCTG      |
| Mn99049  | AACATGCAGTACGAGGAACTCAACAACAAGGGACC---GTGGGACATTTTTCTGGGCGGCTG      |
| Mn11037  | AACATGCAGTACGAGGAACTCAACAACAAGGGACC---GTGGGACATTTTTCTGGGCGGCTG      |
| Mn10106  | AACATGCAGTACGAGGAACTCAACAACAAGGGACC---GTGGGACATTTTTCTGGGCGGCTG      |

| Mb07020  | AACATGCAGTACGAGGAACTCAACAACAAGGGACC---GTGGGACATTTTTCTGGGCGGCTG      |

| Mn12262  | CACTGCACCGGCAGCGGGTGCAGCAG---CCTCAGACCCCCCGGGCAACATCGATGTGCTG       |
| Mn10095  | CACTGCACCGGCAGCGGGTGCAGCAG---CCTCAGACCCCCCGGGCAACATCGATGTGCTG       |
| Mn99049  | CACTGCACCGGCAGCGGGTGCAGCAG---CCTCAGACCCCCCGGGCAACATCGATGTGCTG       |
| Mn11037  | CACTGCACCGGCAGCGGGTGCAGCAG---CCTCAGACCCCCCGGGCAACATCGATGTGCTG       |
| Mn10106  | CACTGCACCGGCAGCGGGTGCAGCAG---CCTCAGACCCCCCGGGCAACATCGATGTGCTG       |

| Mb07020  | CACTGCACCGGCAGCGGGTGCAGCAG---CCTCAGACCCCCCGGGCAACATCGATGTGCTG       |

| Mn12262  | AGGCTCAAGCTCCGCGCCAATCCCCTGATGAGCTGACCTGACCTGGAAGAAGGGCAAGGGGGCTG   |
| Mn10095  | AGGCTCAAGCTCCGCGCCAATCCCCTGATGAGCTGACCTGACCTGGAAGAAGGGCAAGGGGGCTG   |
| Mn99049  | AGGCTCAAGCTCCGCGCCAATCCCCTGATGAGCTGACCTGACCTGGAAGAAGGGCAAGGGGGCTG   |
| Mn11037  | AGGCTCAAGCTCCGCGCCAATCCCCTGATGAGCTGACCTGACCTGGAAGAAGGGCAAGGGGGCTG   |
| Mn10106  | AGGCTCAAGCTCCGCGCCAATCCCCTGATGAGCTGACCTGACCTGGAAGAAGGGCAAGGGGGCTG   |

| Mb07020  | AGGCTCAAGCTCCGCGCCAATCCCCTGATGAGCTGACCTGACCTGGAAGAAGGGCAAGGGGGCTG   |

| Mn12262  | TCGATGATTGGGCGCCGCAGGCAACAAATCGGACACTGACCTGATGACCTGGAAGAAGGGCAAGGGGGCTG |
| Mn10095  | TCGATGATTGGGCGCCGCAGGCAACAAATCGGACACTGACCTGATGACCTGGAAGAAGGGCAAGGGGGCTG |

201
TGGACGACGTGCCGACAC-TATTGACGATCTACTGCTGAAGGCTGACACGACGTGCCC
TG--CGGGACGACGGGCTGCCCAATCCCTGCTGCTGGAAGGCTGACACGACGTCGCCC
TGGACGACGTGCCGACAC-TACTGCAAGCCTCCGCTCTGAGAAGCAGGCTGCCC
* * * * * * * * *
Mn11037  TGCCGACCTCC--TCTTGACGCGCCGAGACCCCAATGACCTCG---------CAG
Mn10106  AGC--TACCCCTGAGACCGCCATCTTCACAGAGAAACGCGCCCATCTGTATCGCCGAG
Mb07020  GGT--GCCACCGGAGTAACAATA--TCCACGGGCGCCCCAGCCCAGTTAATGTGACAGG

Mn12262  GTG--GCCGC--GGTAGCTTGATGTTGCAACTTCGACAGAAGAGC--TACAGACAGGAGACG
Mm10095  CTAACHTCAG--GTTGAAGGAAGAGCAACTTACGAGCTTGCCGAGT--GATGTC
Mm99049  CTAATACAC--GTTGAAGGAAGAGCAACTTACGAGCTTGCCGAGT--GATGTC
Mn11037  GCACTTCGCC--GCACCCCGCAGTCTCGACCTCTCTCTCTCAG---------------GAG--ATGCGACAT-------
Mn10106  GCGGCCGCCCACTCTTTTATAGCTGGCCTCGAGTTCATGGAGTGA-------GCTCCCCCC
Mb07020  CGAAGGCG--ACTGGAAGAGAGGACATCCCGGACTGCGATGGA-TATGTC

Mn12262  GCGGACCGGAGCGGACGGGATCACGCCCC--TCGTGGCGCTACTG--------GGGCCTGCTGC
Mm10095  GCCCAACCCGACGAGCGGCACATCT-CGTTGGGCTCTTCG--GGCTCGGCCATCTGCTGC
Mm99049  GCCCAACCGGAGCGGACGCGGCACATCT--CGTTGGGCTCTTCG--GGCTCGGCCATCTGCTGC
Mn11037  -------ATCGTTGCCGGAAGTTGCTTGGCCTCTGCTG--CTATGCGAGATTTGGC--
Mn10106  GCTCTCTGGGAACCCGAGAACGAACAACTGCTTATTGAGTGGCAGCGCAACAGACGAC
Mb07020  ACAGGCGACGGGGATCG--GATCTTCCTGGGGTTTCTCG--AGCATGCACCAGTTCGCCG

Mn12262  TGGCGCCAGCGCCGAAAACAAGATCATCAGTATCCCCGACAAAGAGCAGCATTACGTCTACCA
Mm10095  TGGCGCCAGCGCCGAAAACAAGATCATCAGTATCCCCGACAAAGAGCAGCATTACGTCTACCA
Mm99049  TGGCGCCAGCGCCGAAAACAAGATCATCAGTATCCCCGACAAAGAGCAGCATTACGTCTACCA
Mn11037  TGGCGGACCGC--AAAAAAAAAG-----GCTCGGACGAGCACAG-----GCTCGGCCGCT
Mn10106  CGACGGCCGGACACTGTGACCGGCTACGCGACCTCGCCGCTGCCTCCAAAAGGCGAGCGGGCTTTACCTTG
Mb07020  GCCGCGGCGCCGACAAAGATCATCAGTATCCCCGACAAAGAGCAGCATTACGTCTACCA

Mn12262  GGGCGACTCCCT--CAGCGTCAGAG--TAGTGACCGCAGCTACCCGAGCC--CTGAGCTGCC
Mm10095  AGGCAGCTCGGT--GACAGAAGAAGACATACATACCCGCACAAAGGAGCAGGACGATCATCTACCCG
Mm99049  AGGCAGCTCGGT--GACAGAAGAAGACATACATACCCGCACAAAGGAGCAGGACGATCATCTACCCG
Mn11037  TGGCGGAGCGC--AAAAAAAAAG-----GCTCGGACGAGCACAG-----GCTCGGCCGCT
Mn10106  GTACGGCAATGAGGAGTGGCAGGACTGCTCCATCGGCGTCATCAACAGCTTACCTTCC
Mb07020  GCCGCGGCGCCGACAAAGATCATCAGTATCCCCGACAAAGAGCAGCATTACGTCTACCA

Mn12262  TGGTCTGCTTTG--AACAGGGCCAGCGGATGTCATTGTCG---------------------
Mm10095  TGGTCTGCTTTG--AACAGGGCCAGCGGATGTCATTGTCG---------------------
Mm99049  TGGTCTGCTTTG--AACAGGGCCAGCGGATGTCATTGTCG---------------------
Mn11037  TGGTCTGCTTTG--AACAGGGCCAGCGGATGTCATTGTCG---------------------
Mn10106  TGGTCTGCTTTG--AACAGGGCCAGCGGATGTCATTGTCG---------------------
Mb07020  TGGTCTGCTTTG--AACAGGGCCAGCGGATGTCATTGTCG---------------------

Mn12262  TGGTCTGCTTTG--AACAGGGCCAGCGGATGTCATTGTCG---------------------
Mm10095  TGGTCTGCTTTG--AACAGGGCCAGCGGATGTCATTGTCG---------------------
Mm99049  TGGTCTGCTTTG--AACAGGGCCAGCGGATGTCATTGTCG---------------------
Mn11037  TGGTCTGCTTTG--AACAGGGCCAGCGGATGTCATTGTCG---------------------
Mn10106  TGGTCTGCTTTG--AACAGGGCCAGCGGATGTCATTGTCG---------------------
Mb07020  TGGTCTGCTTTG--AACAGGGCCAGCGGATGTCATTGTCG---------------------
Chapter 4 Population Genetics

4.1 Introduction

4.1.1 Genetic diversity

Population genetics seeks to describe the fluctuation in the genetic makeup of a population over time (Hartl and Clark 2007). This change in the relative frequencies of different alleles among the population is influenced by factors that include the size of the population, the mating frequency, and the mutation rate. Broadly, a population with a large number of different alleles has higher genetic diversity relative to a population consisting of genetically-similar individuals. Among pathogens, genetic diversity represents an important reservoir of tools that may help a pathogen to attack a wider range of hosts, to resist certain fungicides or plant defenses, and to survive in diverse conditions (Agrios 2005). Although genetic diversity among a species as a whole may be important for the long-term persistence of the pathogen, within a particular population of an agricultural pest (e.g. within a single field), genetic diversity may be restricted due to strong selective pressure from fungicides or from plants that have been bred for resistance (Zhan et al. 2003).

In a large, randomly-mating population with low levels of mutation, selection pressure, and incoming gene flow, the assortment of alleles in a generation of offspring is random, based only on the frequencies of each allele in the parent population (Hartl and Clark 2007). This is known as the Hardy-Weinberg (HW) principle (Hartl and Clark 2007). This model may be useful in predicting gene flow within a population. However, the assumptions inherent in the HW principle may be violated in real populations for several reasons: for example, loci that are physically proximate tend to move together during phenomena such as crossing-over, which
shuffle the parental alleles. The relative frequencies of two such alleles in an offspring population are non-independent. Similarly, major changes in genome organization such as inversions prevent recombination within the inverted segment during meiosis (Hartl and Clark 2007). Linkage disequilibrium occurs when alleles are inherited non-randomly for these or other reasons (Hartl and Clark 2007), and thus can occur within individuals or populations that are subjected to HW conditions in other respects.

Another cause of linkage disequilibrium is a reduced rate of sexual reproduction and/or an increased frequency of asexual reproduction (Hartl and Clark 2007). Clonal reproduction may preclude recombination and can result in the inheritance of a full suite of alleles as a single unit. Although large asexual populations with a high mutation rate may have allele frequencies similar to those present in populations with sexual recombination (Chasnov 2000), small populations that reproduce primarily or solely asexually violate the assumptions of HW equilibrium because the decreased frequency (or absence) of recombination causes allele re-assortment to become non-random. As a result, linkage disequilibrium may be present in such populations.

4.1.2 Mechanisms of generating genetic diversity

In fungi, sexual reproduction (Chapter 5), horizontal gene transfer (Mehrabi et al. 2011), and random mutation (Agrios 2005) may all produce novel combinations of alleles in each generation. The latter two processes are particularly important among fungi that rarely (or perhaps never) reproduce sexually (Webster and Weber 2007). Horizontal gene transfer occurs when the somatic hyphae of separate strains fuse and exchange nuclei (Saupe 2000). The DNA from these nuclei may undergo recombination, thus producing offspring with novel genetic combinations in the absence of meiosis or a sexual cycle (Webster and Weber 2007). The ability
of two strains to participate in horizontal gene transfer is dependent upon their vegetative compatibility group (VCG) identity, which in turn is defined by the alleles present at several loci in the genome (Saupe 2000). Regardless of whether they are compatible, when two hyphae meet, anastomosis (cell fusion) occurs. In the case of compatible hyphae (belonging to the same VCG), nuclei may be exchanged but in incompatible pairings, the fused cells die and no exchange occurs (Saupe 2000). Compatible pairings (resulting in nuclear exchange) have been reported to occur even across species (Friesen et al. 2006), and this form of genetic exchange has been proposed as the mechanism that was responsible for the transformation of the previously non-virulent *Stagonospora nodorum* into a serious pathogen of wheat (Friesen et al. 2006).

The movement of transposable elements (or transposons) is another mechanism that may generate novel combinations of sequence elements without sexual recombination. Transposable elements are fragments of DNA that can change their position within a genome (Hartl and Clark 2007). Although widely present in the genomes of both prokaryotes (Hartl 1992) and eukaryotes, including fungi (e.g. (Parlange et al. 2011; Torriani et al. 2011)), there is also evidence that transposons can move between organisms, including between distantly related species (Renner and Bellot 2012; Schaack et al. 2010). In this manner, transposons may transfer genetic information both within and between species. In addition, the mechanism of insertion and excision employed by some groups of transposons is imprecise and may result in the transfer of genetic information that was not originally part of the transposon (Yu et al. 2000). Transposition is thus another mechanism through which horizontal gene transfer may occur. The contribution of transposable elements to genetic diversity and pathogenicity is further described in Chapter 3.
4.1.3 Repetitive DNA sequences and genetic diversity

To assess the level of genetic variation within a population of fungi, several different tools are available. Repetitive sequences, which are present in varying numbers and various lengths, are one important tool that have been used by researchers (Jarne and Lagoda 1996).

After the double-helix structure of DNA was elucidated, researchers discovered that, under conditions that mimic the physiologic environment, double-stranded DNA most commonly assumes a conformation known as B-DNA consisting of a right-handed double helix with 10.5 nucleotides per turn of the helix (Wang 1979). However, other conformations, including a left-handed helix with 12 nucleotides per turn (Z-DNA), have also been recognized for over 30 years (Rich et al. 1984). Researchers observed that Z-DNA was associated with repetitive sequences consisting of alternating pyrimidines and purines (or vice-versa) (Arnott et al. 1980), and short spans of the repeated pattern G-T were detected in Homo sapiens (Hamada and Kakunaga 1982), Mus musculus (mice) (Nishioka and Leder 1980), Drosophila sp. (Nordheim et al. 1981), and Bos taurus (calf) and Oncorhynchus sp. (salmon) (Hamada and Kakunaga 1982). These simple sequence repeats (SSRs), also commonly known as microsatellites or short tandem repeats (Tautz 1993) are widely found in genomic DNA, including in both coding and in non-coding regions (Katti et al. 2001). These regions consist of a simple pattern of 1-6 nucleotides, repeated between five to approximately one hundred times at each locus, with several thousand such loci found within a single genome (Tautz 1993). The length of these regions is most likely modulated by errors in DNA replication, including both the "slippage" of DNA polymerases (Levinson and Gutman 1987) and errors in the normal error correction mechanisms (Strand et al. 1993). The large number of such loci in eukaryotic genomes, in addition to the variable length of each individual repeat of a given sequence, makes
SSRs an ideal tool for studying genetic diversity. The amplification of a particular SSR from a single individual would be expected to produce multiple bands (corresponding to the multiple loci), each of differing length (reflecting the variable number of repeats at each locus) (Hartl and Clark 2007).

Inter-Simple Sequence Repeats (ISSRs) are the regions located between SSRs (Figure 4.1). Whereas the study of SSRs requires prior knowledge of the genetic sequence of the regions flanking the repetitive sequences (Jarne and Lagoda 1996), the primers to amplify ISSRs are anchored in the more widely-conserved repetitive sequences (McGregor et al. 2000). Because the number of SSRs is the genome is highly variable, the number and the length of the sequences between them are also variable (McGregor et al. 2000). Together, these factors make ISSR analysis useful for population studies even among groups where prior genomic information is unavailable (McGregor et al. 2000).

Analyses using both SSRs and ISSRs have been extensively used to assess genetic diversity in plants (e.g. (Fernández et al. 2002; Galván et al. 2003; Nybom 2004; Patzk 2001; Reddy et al. 2002)) and fungi (e.g. (Karaoglu et al. 2004; Mishra et al. 2003; Wang et al. 2005). Isolates of the plant pathogenic *Fusarium culmorum* were assessed using ISSR (Mishra et al. 2003). Similar to *M. nivale*, this pathogen had been classified as asexual, but was known to possess a high level of phenotypic variation. Investigation of ISSR banding patterns for isolates collected from Europe and North America led these researchers to propose that the geographic populations of this fungus are distinct, and that the variability observed is consistent with recombination (Mishra et al. 2003). Similarly, ISSR analysis revealed genetic variation between isolates of *Beauveria bassiana* and three other *Beauveria* spp. collected from different geographic regions, but not from different host plant species (Wang et al. 2005).
When ISSRs were used to investigate differences among isolates of *Fusarium poae* collected in Argentina and England, the researchers identified a high level of genetic diversity even among isolates collected from the same batch of seeds (Dinolfo et al. 2010). In addition, the isolates did not resolve into two groups based on their country of origin; instead, sub-groups consisting of isolates from a single geographic origin were scattered throughout the dendogram constructed to visualize the results. Even among isolates of the asexual bean pathogen *Pseudocercospora griseola* collected from a single population, ISSR analysis identified a unique haplotype for each of the isolates tested (Abadio et al. 2012).

The analysis of ISSR has been used to assess the contribution of recombination to genetic diversity within a small population of phytopathogenic fungi with a known sexual cycle. In a recent investigation of *Rhizoctonia solani* isolates collected from a single field, Zheng and colleagues suggested that this population likely spreads asexually based on the calculated index of association (2013). The use of markers related to microsatellites provides an economical tool that allows researchers to investigate the genetic diversity among fungal populations.

### 4.1.4 Genetic diversity in *M. nivale*

Outside of the genetic, morphological, and pathogenic differences reported between *M. nivale* and *M. majus* (Chapters 2 and 3), several studies have demonstrated variation within *M. nivale* as well. A study comparing isolates of *M. majus* and *M. nivale* originally collected on European wheat found that, relative to *M. majus*, the RAPD profiles of *M. nivale* were highly diverse (Lees et al. 1995). Maurin and colleagues examined the ITS RFLP patterns among isolates of *M. nivale* sensu lato collected from several host species and from across Europe (1995). This comparison revealed two distinct groups that corresponded to *M. nivale* and *M.
majus. Protein electrophoresis revealed diverse esterase profiles in all twenty of the *M. nivale* isolates examined in this study; by comparison, all seven of the *M. majus* isolates examined possessed the same esterase profile (Maurin et al. 1995). Similarly, in a study comparing the EF-1α sequences of *M. nivale* to those of *M. majus*, researchers noted a slightly higher level of heterogeneity among the *M. nivale* sequences than those of *M. majus* (Glynn et al. 2005). The wider host range of *M. nivale* relative to *M. majus* (Simpson et al. 2000) may provide a possible explanation for the higher variability reported for this species. A study conducted in 1998 (Mahuku et al.) investigated isolates of *M. nivale* collected from four sites consisting of three different turf grass hosts. Using RAPD and RFLP, a high level of genetic variation was observed. The two populations collected from the same host species displayed higher similarity to each other than to either of the other populations despite being collected about 20 km apart.

To explain the high level of genetic variation in *M. nivale* relative to the closely-related *M. majus* (Chapters 2, 3), some researchers have suggested that *M. nivale* may undergo sexual reproduction more frequently than *M. majus* (Mahuku et al. 1998). However, formation of perithecia has been observed more frequently in *M. majus* than in *M. nivale* (Lees et al. 1995). Although the studies described above have reported genetic variation of isolates from several different locations, none have reported re-sampling from the same location across multiple years to determine how rapidly, if at all, genetic variation may fluctuate within a single population. A high turnover in genotype frequency would suggest a sexually recombining population, whereas a high overall level of diversity that is relatively unchanged from year to year could instead suggest an established population of variable individuals that reproduce predominantly or exclusively asexually.
A detailed investigation of the factors that may influence the presence of Hardy-Weinberg equilibrium within *Microdochium nivale* and *M. majus* has not been conducted. However, selective pressure imposed by fungicide use against *M. nivale* on turfgrasses and the apparent host specificity exhibited by both species (Chapter 3) may play an important role in their genetic diversity.

**4.1.5 Objectives**

The objective of this project was to assess genetic variation in local populations of *M. nivale* separated temporally and spatially, with the intention of using these results to assess whether sexual reproduction may be occurring within the populations studied.

**4.2 Materials and Methods**

**4.2.1 Sample collection and DNA extraction**

Immediately following snowmelt in 2011, 2012, and 2013, samples of grass displaying symptoms of pink snow mold and/or Fusarium patch (Figure 4.2) were collected from various grasses at the Guelph Turfgrass Institute (GTI). The sample collection locations are indicated in Figure 4.3, and a list of the total number of samples obtained at each location is found in Table 4.1. The grass samples collected were processed as described in Section 2.2.1 to obtain single isolates per sample. Genomic DNA was extracted from all of the isolates using the modified Edwards protocol (Edwards et al. 1991) as described in Section 2.2.2, and the quality and quantity of the DNA obtained was assessed by electrophoresis as described Section 2.2.2. The DNA was diluted as necessary with dH$_2$O to achieve a concentration of 0.1 ng/µL prior to use in PCR.
4.2.2 SSR primer design

Primers for SSR were designed using the default settings of the program QDD (Meglécz et al. 2010) and using the genome of *M. nivale* isolate 11037 (Chapter 3). Twelve primer sets with penalty values below 7 and a difference in Tm of less than 0.5 °C were selected randomly for synthesis and testing. The SSR primers were synthesized by BioBasic (Markham, Ontario), and all of the primer sets ordered were screened using a PCR protocol identical to that described for the β-tubulin PCR reactions (Chapter 2), with the exception that the annealing temperature was first tested at 50 °C and was raised as necessary to reduce streaking. Primers used in the study are listed in Table 4.2.

4.2.3 ISSR and SSR PCR protocols

All ISSR and SSR PCRs were conducted using the same BioRad MyCycler thermocycler (BioRad, Hercules, California). The ISSR primers were obtained from either the local lab collection of ISSR primers or from the literature. All ISSR primers were synthesized by Laboratory Services, University of Guelph (Guelph, Ontario). A list of the primers tested is found in Table 4.2. Each ISSR PCR reaction contained approximately 0.1 ng of fungal DNA (0.1 ng/μL) with a final concentration of 1.2X PCR buffer, 0.22 mM Mg²⁺ (BioBasic, Markham, ON), 1.5 mM dNTP mixture (prepared using individual solutions of dATP, dTTP, dCTP, and dGTP; BioBasic, Markham, ON), 0.6 μM of each the forward and reverse primers, 0.04 U of *Tsg* DNA polymerase (BioBasic, Markham, Ontario), and enough sterile water to bring the total volume of the master mix to 20 μL per reaction.

The ISSR primers were screened by testing the DNA from ten samples collected in 2011 (five each from the "roadway" and the "pathology green fringe" locations). The reaction mixture
was submitted to a thermal cycling procedure consisting of an initial denaturation at 94 °C for 1.5 minutes followed by 35 repetitions of a denaturation period at 94 °C for 40 seconds, an annealing period of 45 seconds at 45 °C, followed by an extension period of 1.5 minutes at 72 °C. The mixture was then subjected to a final extension period of 5 minutes at 72 °C before being held at 15 °C until the sample was retrieved and stored in the refrigerator at 4 °C.

Those primers that yielded at least one band that was present in at least one, but not all of the isolates tested, were selected for further screening. The annealing temperatures for these primers were optimized using a gradient PCR. The temperatures included in the gradient were 42.0, 42.7, 44.0, 46.0, 48.6, 50.7, 52.1, and 53.0 °C, and higher annealing temperatures were tested if streaking was observed. The annealing temperatures for the primers selected for use with all isolates are found in Table 4.3.

For SSR, the PCR reaction mixture was prepared as described for the ISSR reactions. Based on their consistent production of polymorphic bands from a selection of isolates from each collection location from 2011 and 2012, five SSR primers were used to amplify DNA isolated from all of the samples included in this study (Table 4.3).

4.2.4 Result scoring and analysis

To visualize the results of the test SSR and ISSR PCRs, 1.2% agarose gels were prepared using 0.5X TBE buffer. A 3:1 mixture of PCR product and loading buffer (Chapter 2) was loaded onto the gel and the samples were subjected to electrophoresis in 0.5X TBE buffer at 50 V for approximately 60 minutes before being stained with an ethidium bromide solution and visualized under UV light. For the PCRs that included all isolates from a given year, large, 72-well 1.2% agarose gels were instead subjected to electrophoresis at 85 V and 35 mW for 4 h.
Commercial DNA ladders (both 100bp and 1kb) (described in Chapter 2) were also loaded in each gel prepared in 2011 and 2012 so that the sizes of the bands could be estimated. For the 2013 samples, the ladders used were a 100 bp ladder (BioBasic, Markham, ON), which hand bands ranging from 0.1 to 1.5 kb in length, and the Lambda DNA / HindIII plus marker (BioBasic, Markham, ON), which had bands ranging from 0.3 to 23 kb in length.

Each gel was photographed electronically and the presence or absence of bands for all of the samples amplified with the eight ISSR and SSR primers or primer sets were scored in a binary matrix. Scoring was performed by eye. The data for all primer sets were combined and were used to construct a UPGMA tree by using the program WinDist (Yap and Nelson 1996) to calculate a distance matrix, and the "neighbor" function of PHYLIP v. 3.69 (Felsenstein 1993) to translate this matrix into a tree. Bootstrap values were calculated using WinBoot (Yap and Nelson 1996). Trees were visualized and saved as image files using Archaeopteryx (Han and Zmasek 2009), and bootstrap values (calculated by WinBoot) were added to the figure in PowerPoint.

Tests for linkage disequilibrium were conducted using two different methods. An exact test for linkage disequilibrium for haplotypic data was conducted using Arlequin v. 3.1 (Excoffier et al. 2005). The second test was conducted using the Index of Association (IA) (Smith et al. 1993) to compare the observed variance in the observed number of mismatches between isolates ($V_{obs}$) with the expected variance ($V_{exp}$) of mismatches based on allelic frequencies for a randomly recombining system without linkage disequilibrium. The IA is calculated as:

$$IA = \frac{V_{obs}}{V_{exp}} - 1$$

(equation 4.1)
If IA is larger or smaller than 0, linkage disequilibrium is present in the system. For this study, if the absolute value of IA was greater than 2, linkage disequilibrium was scored as present. To assess the variation in IA, 100 bootstrap replicates were conducted to generate data sets with the same number of isolate and allele frequencies for each locus but with random assignments of alleles to isolates without the assumption of linkage (Feil et al. 1996). This test (the calculation of IA and bootstrap testing) was implemented using the program DISEQUIL (Mahuku et al. 1998; program available at: www.uoguelph.ca/~thsiang/pubs/supplement/disequilibrium/disequil.zip). Using both methods, linkage disequilibrium was assessed separately within each "by-year" group, "location" group, and "year and location" group. An example of the input file for Arlequin is found in Appendix 4.1. For fully clonal isolates (i.e. groups of isolates with identical banding patterns across all eight primers), only one representative was included in the input for these calculations.

4.3 Results

4.3.1 ISSR primer screening

A total of 12 ISSR primers previously used in the lab and from the literature were obtained or synthesized and were tested with five isolates for each of the two locations from the 2011 collection year. Eleven of these primers produced at least a single band with each isolate, and the PCR conditions were further optimized by performing a gradient PCR. When the primers were tested with the samples from 2012 and 2013, only three of the eleven primers produced polymorphic bands. For this reason, only the primers BHY(AGC)$_5$, (GA)$_8$T, and DD(CCA)$_5$ were used to amplify all of the DNA samples. These primers yielded between zero to eight bands
per isolate (Appendix 4.2). One of the primers ((CT)$_3$RG) did not amplify any of the isolates tested.

The primer BHY(AGC)$_5$ produced 12 polymorphic bands, ranging in size between 0.4 to 2.1 kb. The number of bands observed for any single isolate ranged between zero to a maximum of eight, with an average of five bands per isolate. The primer (GA)$_8$T produced 19 polymorphic bands, ranging in size between 0.26 to 2.2 kb. The number of bands observed for any single isolate ranged between zero to six, with an average of three bands per isolate. The primer DD(CCA)$_5$ produced 19 polymorphic bands, ranging in size between 0.4 to 2.2 kb in size. The number of bands observed for any single isolate ranged between zero and eight, with an average of five bands per isolate.

4.3.2 SSR primer design and screening

Using the program QDD, a total of 41,770 primer sets were designed. Twelve of the primer sets were ordered and were tested with a subset of five isolates from each of the two locations and from both the 2011 and 2012 collection years (i.e. a total of 20 isolates). Five of these primer sets produced at least one polymorphic band and were used to amplify the DNA from all of the isolates. The primers used to amplify all of the isolates were 961061, 910478, (CT)$_5$, (GAT)$_6$, and (GC)$_5$, and these yielded between zero to eight bands per isolate (Appendix 4.2).

Primer pair 961061 produced 16 polymorphic bands, ranging in size between 0.17 to 3.0 kb in length. The number of bands observed for any single isolate ranged between zero and eight, with an average of four bands per isolate. Primer pair 910478 produced 17 polymorphic bands, ranging in size between 0.07 to 5.0 kb in length. The number of bands observed for any single
isolate ranged between zero and ten, with an average of four bands per isolate. Primer pair (CT)$_5$ produced 23 polymorphic bands, ranging in size between 0.09 to 2.5 kb in size. The number of bands observed for any single isolate ranged between zero and nine, with an average of four bands per isolate. Primer pair (GAT)$_6$ produced 10 polymorphic bands, ranging in size between 0.2 to 1.3 kb in size. The number of bands observed for any single isolate ranged between zero and five, with an average of two bands per isolate. Primer pair (GC)$_5$ produced 20 polymorphic bands, ranging in size between 0.075 to 5.0 kb in size. The number of bands observed for any single isolate ranged between zero and nine, with an average of three bands per isolate.

When the isolates were compared using all of the polymorphic SSR and ISSR bands produced, only two clonal isolates (collected from the pathology green fringe in 2011) were identified. However, within the results for each primer, several isolates shared identical banding patterns. A table summarizing these relationships for each primer is found in Appendix 4.2.

**4.3.3 Linkage disequilibrium calculations**

In the linkage disequilibrium calculations performed by Disequil (Table 4.4), linkage disequilibrium was detected among all of the isolate groups collected at each location and in each of the three years, with the exception of the six samples collected near the pathology green in 2011. Linkage disequilibrium was also calculated in Arlequin, using a p value of 0.05. By default, Arlequin assesses only those alleles which are polymorphic within each of the groups defined.

In the 2011 samples, linkage disequilibrium was not found among any of the 26 polymorphic loci from the six pathology green samples, but was detected in 20 of the 37 polymorphic loci from the roadside samples. For the 46 samples from 2012 and the 74 samples
from 2013, linkage disequilibrium was detected among the polymorphic loci of all samples (Table 4.4).

### 4.3.4 Year-to-year and location-to-location comparisons

The bootstrapped UPGMA dendogram (Figure 4.4) revealed some trends among the yearly isolate groups from isolates collected in different years. For the 2013 collection, 47 of the 74 total isolates were grouped together in 50% of the trees produced. An additional 25 isolates formed a separate group in 32% of the trees. The remaining two isolates were grouped together with 100% bootstrap support. Among the 2012 isolates, 45 of the 46 isolates formed a single group with 58% bootstrap support. The remaining 2012 isolate was external to the group containing the other 2012 isolates, but in 29% of the trees produced, it was present in a sister group relative to the group containing the larger "2012" group. In 95% of the trees, all of the 2011 isolates were grouped together. In addition to these within-year trends, the 2013 isolates were split into two large groups of unequal size as described above and one much smaller group consisting of just two isolates. The largest group was a sister group to the 2011 group in 39% of the trees examined. The 2013 two-isolate group was a sister group to the 2012 isolates in 14% of the trees examined, and the final group of 2013 isolates were clustered with the larger group that contained both the 2012 isolates and the 2013 isolate pair in 20% of the trees sampled.

As a whole, the isolates collected from the two locations did not group together, but there were some trends for the samples collected at the two locations within each year. Within the 2013 isolates, 28 of the 43 "pathology green" isolates formed a paraphyletic group with 23% bootstrap support; an additional 6 isolates formed a single group with 98% bootstrap support. Among the 31 "roadside" isolates collected in 2013, two separate groups of four isolates formed
distinct groups, each with 100% bootstrap support; an additional seven 2013 "roadside" isolates were grouped together with 77% support. The remaining isolates from both locations formed polyphyletic groups.

For the isolates collected in 2012, all but one of the 17 "roadside" samples formed a group with 48% bootstrap support. The 25 "pathology green" 2012 isolates formed a single group with 54% bootstrap support. Among the 2011 isolates, three of the five "pathology green" isolates formed a single group with 22% bootstrap support. There were no strongly-supported (i.e. bootstrap values above 50%) non-polyphyletic groups containing the nine 2011 "roadside" isolates.

4.4 Discussion

In this Chapter, the genetic diversity of two proximate populations of *M. nivale* was assessed yearly during a three-year period. The goal of these experiments was to determine whether genotypes persisted from year to year and to explore whether recombination may be occurring in these populations. The isolates that were examined in these experiments were collected from two proximate patches of Kentucky bluegrass at the Guelph Turfgrass Institute, immediately following snowmelt in 2011, 2012, and 2013.

At the beginning of the study, the first tools that were used to assess the genetic diversity among these samples were ISSR markers. A set of 12 primers were tested, but only three ultimately yielded polymorphic bands for the isolates under study. One primer failed to amplify any of the isolates tested while the other eight primers did not yield polymorphic bands (but did amplify at least one band from the isolates tested). Among the three primers that did yield
polymorphic results, either 12 or 19 polymorphic bands were produced. These bands were scored as polymorphic because they were observed in at least one, but not all of the isolates tested.

Because only three ISSR primers yielded polymorphic bands, SSR primers were designed using the genome of *M. nivale* isolate 11037 and the program QDD to further explore the genetic diversity within the populations of interest. The program identified over 46,000 possible SSR primer pairs from the *M. nivale* genome, which is consistent with the results obtained from the similarly-sized genomes of other Ascomycetes (Fang Shi and Mihaela Stanescu, personal communication). Although other researchers have used QDD to detect SNPs in fungi (e.g. (Rouxel et al. 2012)), the number of putative SSR primers detected was not reported.

Of the 12 SSR primer sets ordered and tested, five were selected for use with all of the isolates collected based on their reproducible production of polymorphic bands. The remaining seven primers amplified at least one band from the isolates they were tested with, but failed to yield polymorphic results. This result is discussed in more detail below.

Previous research has reported the presence of high levels of genetic diversity among *M. nivale*, which has lead some researchers to suggest that sexual reproduction may be common within this species (e.g. (Mahuku et al. 1998)). A high level of genetic variation was also detected in this experiment, as a total of 130 polymorphic bands were produced by the eight ISSR and SSR primers used in this study. In each year, isolates of *M. nivale* were collected from two locations. Although there was no overall pattern in the distribution of genotypes between the two locations, trends were detected both between and among the three years studied.

The isolate collections from two of the three years of this study formed single groups; the 2013 data, which split into two large, paraphyletic groups and one very small group consisting of just two isolates, was the exception to this observation. The two large groups of 2013 isolates
were grouped more strongly with the 2011 and the 2012 isolates than they did with each other, suggesting that there were distinct populations within the isolates from this collection year. This variability in the 2013 population was independent of collection location, as both of the large groups contained isolates from the two locations.

In addition to the differences observed between the years, within each year, the isolates originally collected from alongside the pathology green formed a separate group relative to the isolates collected from the roadside. This consistent pattern suggests that, despite the proximity of these two locations, there was restricted gene flow between these populations. In addition, the grouping of the isolates by year, and then by location (and not by location and then by year), implies that the year-to-year variability was higher than the location-to-location variability. The splitting of the 2013 isolates into two large groups that were clustered with the 2011 and the 2012 isolates, respectively, implies that new genetic information was introduced into the population of isolates between 2012 and 2013; however, it is difficult to extend this claim to the 2011 due to the small sample size. The low level of clonality detected also supports the hypothesis that asexual reproduction alone does not account for the variability observed. The most likely sources for new genetic information in these populations are sexual reproduction (i.e. novel combinations of alleles that were already in existence in the population), the infestation of the locations of interest with inoculum from an outside source (e.g. from neighbouring fields), or some combination of these two effects.

Although the overall level of clonality among all of the isolates studied was very low, within each primer or primer set there were several isolates which produced identical banding patterns (Appendix 4.2). The identical banding patterns were primarily found within groups of isolates from the same year and collection location, and suggest that, although not fully clonal,
large numbers of isolates did share "islands" of genetic information. These groups of shared alleles may also account for the linkage disequilibrium detected among most of the populations studied. Linkage disequilibrium was detected among all of the year-by-location groups with the exception of the "pathology green" samples from 2011. This non-random inheritance pattern could be caused by several factors including the physical proximity of the loci chosen for study or a lack of recombination among the population.

The pockets of similarity represented by these "islands" of identical banding patterns may explain why linkage disequilibrium was detected among each of the populations, because the banding patterns that were observed among more than three isolates were confined within a single year, and most often within a single collection location. The persistent grouping of certain alleles responsible for any particular banding pattern may imply that isolates sharing this pattern share a common parent; however, a hypothetical parent was not detected among the previous years' isolates that were included in this study.

The apparent linkage of many loci within the individuals tested also agrees with some of the ISSR and SSR results obtained during the primer testing phase. For both methods, the majority of the primers or primer sets tested (a total of 16 out of 24) failed to yield polymorphic bands. Instead, either identical banding patterns (15 of the 16 "rejected" primers) or no banding at all (one primer) were produced.

Taken together with the results from the primers that did yield polymorphic bands, the results from the non-polymorphic bands suggest that, although there was genetic variation within the populations studied, large portions of the genomes of the isolates sampled were not highly variable. The presence of a high degree of similarity, reflected through both the patterns of similarity within the "polymorphic" primers and through the identical results for all isolates with
the non-polymorphic primers, is broadly consistent with a largely asexual population that experiences infrequent sexual reproduction. Despite the high level of genetic variation reported within local populations of *M. nivale* (e.g. (Mahuku et al. 1998) and this study), the frequency of sexual reproduction within this species is still uncertain (Chapter 5). Although perithecia have been observed *in vitro* for this species, they have not been reported from the field (Smith 1983), suggesting that if *M. nivale* does undergo sexual reproduction, this process may be infrequent or may occur only within a narrow set of conditions (Smith 1983; Tronsmo et al. 2001). Even fungi that are not known to possess a sexual cycle can exhibit a high level of genotypic variation by ISSR (Abadio et al. 2012); regardless, the "pockets" of genotypic similarity observed within each year in this study is similar to the overall trend observed between countries for populations of the homothallic pathogen *F. poae* studied by Dinolfo and colleagues (2010).

To further explore the trends observed within these experiments, more thorough sampling of both the two locations included in this study and the surrounding areas is necessary. A more detailed comparison between these isolates and those collected from the surrounding fields, which consist of a different turfgrass species, may elucidate possible sources for new genetic information. In addition, the presence of horizontal gene transfer in *M. nivale* was not examined, and may prove to be an important mechanism of gene transfer within this population.

The results presented in this chapter show that the genotypes present within two small populations of *M. nivale* were variable when sampled yearly across a period of three years using SSR and ISSR markers. However, this variability may have been limited to only small portions of the genome, because linkage disequilibrium calculations and examinations of the banding patterns within primers or primer sets suggested that, despite possessing some differences, isolates that were ultimately ranked as possessing a unique genotype actually contained several
loci that were identical to those of other individuals within their year and/or location collection group. These results underscore the necessity of examining a large number of isolates for a study of this kind, and suggest that a low level of sexual reproduction may be occurring in *M. nivale* in the field. The ability of *M. nivale* to undergo sexual reproduction is further explored in Chapter 5, which also includes an investigation into the mating type sequences in both *M. nivale* and *M. majus*. 
4.5 References for Chapter 4


Table 4.1 Year and location of collection from the Guelph Turfgrass Institute for all samples included in multi-year screening. See Figure 4.3 for a map depicting these locations.

<table>
<thead>
<tr>
<th>Year</th>
<th>Collection location and grass species</th>
<th>Near Pathology Green (PG)</th>
<th>Near Roadside (Rd)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td></td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>2012</td>
<td></td>
<td>25</td>
<td>21</td>
</tr>
<tr>
<td>2013</td>
<td></td>
<td>43</td>
<td>31</td>
</tr>
</tbody>
</table>
**Table 4.2** List of all SSR and ISSR primers screened to assess genetic variation in *Microdochium nivale* field isolates collected across three years.

<table>
<thead>
<tr>
<th>Primer Type</th>
<th>Primer Name</th>
<th>Primer Sequence(s)</th>
<th>Primer Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ISSR</strong></td>
<td>(ACC)$_6$ CC</td>
<td>ACCACCACCACCACCACCACCC</td>
<td>(Baysal et al. 2010)</td>
</tr>
<tr>
<td></td>
<td>(AG)$_8$</td>
<td>AGAGAGAGAGAGAGAGAG</td>
<td>(Bayraktar et al. 2008)</td>
</tr>
<tr>
<td></td>
<td>(CT)$_5$ RG</td>
<td>CTCTCTCTCTRG</td>
<td>Hsiang Lab</td>
</tr>
<tr>
<td></td>
<td>CT(GA)$_8$</td>
<td>CTGAGAGAGAGAGAGAGAGAGA</td>
<td>(Dinolfo et al. 2010)</td>
</tr>
<tr>
<td></td>
<td>(CCA)$_5$</td>
<td>CCACCCACCCACCA</td>
<td>(Bornet and Branchard 2001)</td>
</tr>
<tr>
<td></td>
<td>(GA)$_8$T</td>
<td>GAGAGAGAGAGAGAGAGAT</td>
<td>Hsiang Lab</td>
</tr>
<tr>
<td></td>
<td>(GA)$_6$GG</td>
<td>GAGAGAGAGAGAGAGAGG</td>
<td>Hsiang Lab</td>
</tr>
<tr>
<td></td>
<td>BHY(AGC)$_5$</td>
<td>BHYAGCAGCAGCAGCAGC</td>
<td>(Schneider et al. 2009)</td>
</tr>
<tr>
<td></td>
<td>CCA(TGA)$_3$TG</td>
<td>CCATGATGATGATGATGATGATG</td>
<td>(Baysal et al. 2010)</td>
</tr>
<tr>
<td></td>
<td>(ACC)$_6$ CC</td>
<td>ACCACCACCACCACCACCA</td>
<td>(Baysal et al. 2010)</td>
</tr>
<tr>
<td></td>
<td>DD(CCA)$_5$</td>
<td>DDCCACCCACCCACCC</td>
<td>(Hantula et al. 1996)</td>
</tr>
<tr>
<td></td>
<td>(CAC)$_5$</td>
<td>CACCAACCCACCCACCA</td>
<td>(Dinolfo et al. 2010)</td>
</tr>
<tr>
<td><strong>SSR</strong></td>
<td>MnSSR_(GC)$_5$ F/R</td>
<td>TGCAGGGACTCATCGACC</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>MnSSR_(CT)$_5$ F/R</td>
<td>TCATCTCCGCACACTCC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MnSSR_(AAC)$_5$ F/R</td>
<td>ACCACCCACCCACCCACCC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MnSSR_(GAT)$_6$ F/R</td>
<td>ACAAGACCGATGACGATGAC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MnSSR_(CG)$_5$ F/R</td>
<td>GTGGAACCTGAGCCGCAC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MnSSR_916061 F/R</td>
<td>CACCAAGCAAAGCGAGGA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MnSSR_908729 F/R</td>
<td>GGCTCGATGCCAGCATA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MnSSR_920718 F/R</td>
<td>GTCATACGGACCCGAG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MnSSR_923044 F/R</td>
<td>GGCTCGATGCCAGCATA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MnSSR_927245 F/R</td>
<td>GAGATGCGCAAGTACCC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MnSSR_920478 F/R</td>
<td>AACTCGCATGGCTGCCTG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MnSSR_908729 F/R</td>
<td>GAGATGCGCAAGTACCC</td>
<td></td>
</tr>
</tbody>
</table>

261
Table 4.3 List of all SSR and ISSR primers selected for analysis of genetic variation in *Microdochium nivale* field isolates collected across three years.

<table>
<thead>
<tr>
<th>Primer Type</th>
<th>Primer Name</th>
<th>Number of Polymorphic Bands</th>
<th>Annealing Temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISSR</td>
<td>BHY(AGC)$_5$</td>
<td>12</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>(GA)$_8$T</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DD(CCA)$_5$</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>SSR</td>
<td>MnSSR_(GC)$_5$ F/R</td>
<td>20</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>MnSSR_(CT)$_5$ F/R</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MnSSR_(GAT)$_6$ F/R</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MnSSR_916061 F/R</td>
<td>16</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>MnSSR_920478 F/R</td>
<td>17</td>
<td></td>
</tr>
</tbody>
</table>
**Table 4.4** Results of linkage disequilibrium calculations performed using the program Disequil (described in Mahuku et al. 1998) on isolates of *Microdochium nivale* collected in three separate years and in two locations at the Guelph Turfgrass Institute. See Figure 4.3 for a map depicting these locations.

<table>
<thead>
<tr>
<th>Collection Year</th>
<th>Collection Location*</th>
<th>Observed variance</th>
<th>Expected variance</th>
<th>Index of association</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>PG</td>
<td>15.067</td>
<td>6.240</td>
<td>1.41</td>
</tr>
<tr>
<td></td>
<td>Rd</td>
<td>32.187</td>
<td>8.0339</td>
<td>3.01</td>
</tr>
<tr>
<td>2012</td>
<td>PG</td>
<td>112.371</td>
<td>15.746</td>
<td>6.14</td>
</tr>
<tr>
<td></td>
<td>Rd</td>
<td>137.890</td>
<td>15.604</td>
<td>7.84</td>
</tr>
<tr>
<td>2013</td>
<td>PG</td>
<td>121.458</td>
<td>15.597</td>
<td>6.79</td>
</tr>
<tr>
<td></td>
<td>Rd</td>
<td>136.514</td>
<td>19.432</td>
<td>6.03</td>
</tr>
</tbody>
</table>

* PG = near pathology green; Rd = along roadside
Figure 4.1 Diagram depicting the relative positions of hypothetical SSR and ISSR loci and the primers that could amplify these regions. The SSR loci (white text on black background) are flanked by ISSR regions (black text on white background). Whereas the ISSR primer (white text on black arrow) is anchored within the repetitive SSR region, and thus can be designed with only knowledge about the repetitive SSR sequence, the SSR primers (black text on white arrows) are located within the non-repetitive ISSR sequences, and thus require more detailed genomic information.
Figure 4.2 Grasses displaying symptoms of pink snow mold and/or Fusarium patch. Both photos were taken at the GTI, Guelph, Ontario, 2 March 2012. a) Kentucky bluegrass in the area to the east of the native green and b) annual bluegrass/creeping bentgrass mixture, native green.
Figure 4.3 Map of the Guelph Turfgrass Institute with collection locations and key landmarks indicated: A: Near pathology green (PG); B: roadside (Rd). The grass species at both collection locations was *Poa pratensis* (Kentucky bluegrass).
Figure 4.4 UPGMA tree depicting relationships between *M. nivale* isolates collected from *P. pratensis* at two locations (Figure 4.3) yearly from 2011-2013. Bootstrap values (out of 100) are displayed on key nodes. Legend: 1: 2013 pathology green isolates; 2: 2013 roadside; 3: 2012 pathology green; 4: 2012 roadside; 5: 2011 pathology green; 6: 2011 roadside.
Appendices for Chapter 4

Appendix 4.1 Sample input for linkage disequilibrium calculation with Arlequin

[Profile]
 NbSamples=6
 DataType=RFLP # - {DNA, RFLP, MICROSAT, STANDARD, FREQUENCY}
 GenotypicData=0 # - {0, 1}
 GameticPhase=1 # - {0, 1}
 LocusSeparator=Tab # - {TAB, WHITESPACE, NONE}
 RecessiveData=0 # - {0, 1}
 MissingData='?' # A single character specifying missing data

# Some advanced settings the experienced user can uncomment
# Frequency= ABS       # - {ABS, REL}
# FrequencyThreshold= 1.0e-5 # - (Any real number, usually between 1.0e-7 and 1.e-3)
# EpsilonValue= 1.0e-7    # - (Any real number, usually between 1.0e-12 and 1.0e-5)

[Data]

[[HaplotypeDefinition]]
 HaplListName="List of observed haplotypes"
 HaplList = ({
 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0
 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0
 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0
 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 2 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0
 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0
 0 0 0 1 0 0 0 1 0 1 1 0 0 0 1 0
 1 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

...}

[[Samples]]

#1
SampleName="2013_pg"
SampleSize= 43
SampleData= {
1  
2  
3  
4  
5  
6  
7  
8  
9  
10 
11 
12 
13 
14 
15 
16 
17 
18 
19 
20 
21 
22 
23 
24 
25 
26 
27 
28 
29 
30 
31 
32 
33 
34 
35 
36 
37 
38 
39 
40 
41 
42 
43 
}

#2
SampleName="2013_rd"
SampleSize= 31
SampleData= {
44 
45 
46 
}
#3
SampleName="2012_pg"
SampleSize=25
SampleData= {
75 1
76 1
77 1
78 1
79 1
80 1
81 1
82 1
83 1
84 1
85 1
86 1
87 1
88 1
89 1
90 1
91 1
92 1
93 1
94 1
}

270
SampleName="2012_rd"
SampleSize=17
SampleData= {
  100 1
  101 1
  102 1
  103 1
  104 1
  105 1
  106 1
  107 1
  108 1
  109 1
  110 1
  111 1
  112 1
  113 1
  114 1
  115 1
  116 1
}

SampleName="2011_pg"
SampleSize=5
SampleData= {
  117 1
  118 1
  119 1
  120 1
  121 1
}

SampleName="2011_rd"
SampleSize=9
SampleData= {
  122 1
  123 1
  124 1
  125 1
  126 1
  127 1
  128 1
  129 1
  130 1
}
[[Structure]]
StructureName="by year"
NbGroups=3
#1
Group=
"2013_pg"
"2013_rd"
}

#2
Group=
"2012_pg"
"2012_rd"
}

#3
Group=
"2011_pg"
"2011_rd"
}
Appendix 4.2 Banding patterns detected with primer groups studied (n= 136). Multiple rows with the same number of bands represent different banding patterns based on band sizes.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Band sizes observed</th>
<th>Number of isolates with a unique banding pattern</th>
<th>Repetitive banding patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Number of bands</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Number of isolates displaying pattern</td>
</tr>
<tr>
<td>MnSSR_916061 F/R</td>
<td>3,000; 1,300; 1,150; 1,000; 850; 700; 525; 475; 450; 425; 400; 325; 300; 250; 170</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>MnSSR_920478 F/R</td>
<td>5,000; 2,000; 1,300; 1,000; 900; 800; 725; 675; 550; 475; 375; 350; 300; 275; 250; 200; 150; 75</td>
<td>86</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>MnSSR_(CT)5 F/R</td>
<td>2,500; 1,400; 1,300; 1,200; 1,000; 950; 900; 800; 750; 625; 600; 525; 450; 375; 300; 260; 225; 200; 160; 100; 90</td>
<td>91</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>MnSSR_(GAT)6 F/R</td>
<td>1,300; 1,250; 1,050; 900; 775; 500; 375; 300; 275;</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Gene</td>
<td>Size (base pairs)</td>
<td>Frequency</td>
<td>Size (base pairs)</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------</td>
<td>-----------</td>
<td>------------------</td>
</tr>
<tr>
<td>MnSSR_(GC)5</td>
<td>5,000; 1,500; 1,100; 1,050; 1,000; 900; 800; 700; 650; 500; 475; 425; 380; 325; 300; 275; 225; 175; 150; 75</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>DD(CCA)5</td>
<td>2,200; 1,900; 1,600; 1,500; 1,300; 1,200; 1,075; 1,050; 1,000; 950; 900; 860; 850; 730; 650; 575; 500; 450; 400</td>
<td>80</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>(GA)8T</td>
<td>2,400; 2,200; 1,700; 1,600; 1,400; 1,200; 1,100; 1,000; 950; 875; 800; 750; 650; 600; 550; 510; 480; 400; 310; 260</td>
<td>113</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>BHY(AGC)5</td>
<td>2,100; 2,000; 1,800; 1,100; 1,050; 900; 800; 700; 620; 550; 480; 400</td>
<td>104</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
</tbody>
</table>
Chapter 5 Mating Type Experiments

5.1 Introduction

5.1.1 Reproduction in Fungi

As a group, fungi, like plants, are capable of both sexual and asexual reproduction. Asexual propagation may occur through the dispersal of mitotically produced spores (conidia) or hyphal fragments, whereas sexual reproduction involves the formation of meiospores (Taylor et al. 1999). Many species of fungi are capable of both mechanisms of reproduction, and may employ both under different conditions (e.g. based on environmental factors, such as temperature or the availability of nutrients). Historically, fungi were classified based on the morphology of their sexual structures (Webster and Weber 2007); however, the sexual cycle of some fungi is completely unknown. In the past, apparently asexual species were placed in the Deuteromycetes, but this taxonomic class has been abandoned in more recent classification systems by re-organizing taxa based on evolutionary relationships established using other tools, such as DNA sequencing data (Hibbett et al. 2007; Schoch et al. 2009). In addition, molecular techniques have facilitated the identification of cryptic sexual reproduction in populations previously believed to be strictly asexual (Arie et al. 2000).

Among species that do reproduce sexually, the mechanisms and requirements of sexual reproduction vary widely. In the Dikarya, which are composed of the Ascomycota and the Basidiomycota, mating is apparently under the control of the genes at a single locus (Butler 2007). In heterothallic Ascomycota, the mating type locus, usually abbreviated as MAT1 (Turgeon and Yoder 2000), contains one of two possible genes (or gene families). While both MAT genes in heterothallic species occupy the same genetic region (i.e. the MAT1 locus) in
different individual isolates, they are referred to as idiomorphs rather than alleles to emphasize their extremely divergent nature (Metzenberg and Glass 1990). Collectively, both genes are referred to as mating type (MAT) genes, and sexual reproduction in heterothallic Ascomycota requires plasmogamy between two individuals that each carry a different idiomorph of the MAT genes.

Although individual isolates of heterothallic Ascomycota possess only a single mating type gene and require an individual with the opposite MAT gene for meiosis to occur, other modes of sexual reproduction are possible. Homothallic species contain both mating type genes within single nuclei, and individual strains (i.e. single-spore- or single-nucleus-derived cultures) are self-fertile and usually do not outcross with any other individual. More complex mating systems also exist, including pseudohomothallism, wherein different nuclei are found in the meiospore, each possessing one of the two MAT genes (Merino et al. 1996). Most ascospores produced in pseudohomothallic species possess both nuclei (and thus contain both MAT genes), but, rarely, individual ascospores may contain only a single nucleus. The cultures derived from these monokaryotic spores are self-sterile but are fertile with cultures derived from either a monokaryotic spore of the opposite mating type, or with cultures containing both nuclei as in the parent tissue (Merino et al. 1996).

Most filamentous Ascomycete species studied may be readily categorized as asexual, homothallic, heterothallic, or pseudohomothallic based on their mating behaviour. Genetic studies examining the MAT genes have generally supported the non-taxonomic groupings based on these observations (e.g. (Yun et al. 2000)), with the intriguing exception that putative MAT genes have been identified in the genomes of apparently asexual species (e.g. (Arie et al. 2000; Yun et al. 2000)). However, unusual mating systems that do not agree with these categories are
known. For example, in the heterothallic plant pathogen *Ophiostoma quercus*, isolates of both mating types are reported to contain both MAT idiomorphs (Wilken et al. 2012).

In a more complex variation from the simple categories described above, several species within the genus *Glomerella* have displayed unusual mating patterns which do not seem to be solely dependent upon the two "canonical" MAT genes in the Ascomycota (Butler 2007). For example, in *G. lindemuthiana*, one cross between self-sterile strains produced asci and ascospores in culture, yet both parental strains apparently contained a single MAT gene (Rodriguez-Guerra et al. 2005). In *G. graminicola*, both self-sterile and homothallic strains have been identified (Vaillancourt et al. 2000), but only a single MAT gene was identified in all isolates tested by PCR (Chen et al. 2002). The single MAT gene that has been amplified in several species of *Glomerella* is a single idiomorph; the opposite idiomorph has not been successfully amplified by PCR (Menat et al. 2012; Rodriguez-Guerra et al. 2005; Vaillancourt et al. 2000) nor identified from the genome sequences available to date (Menat et al. 2012).

Unbalanced heterothallism, wherein individual strains of a species contain mutated fertility genes that may either be compatible or incompatible, was proposed by Wheeler (1954) to explain his observations with this genus; however, molecular evidence for this hypothesis has not been found (Menat et al. 2012). Similarly, among the well-studied genus *Neurospora*, only a single MAT gene has been detected among the species *N. africana*, *N. lineolata*, and *N. galapagoensis*, despite their apparent homothallism (Lin and Heitman 2007). Together, these observations suggest that the relatively simple model of fungal mating as solely under the control of the MAT1 locus may be incomplete for at least some taxa of Sordariomycetes.
5.1.2 Genes associated with sexual reproduction

The current model of sexual reproduction in the Ascomycota states that, as a group, these fungi possess two different mating types, usually referred to as + and −, A or a, or 1 and 2 (Casselton 2008). The names MAT1-1 and MAT1-2 have been suggested as standardized names for these two idiomorphs (Turgeon and Yoder 2000). Each of the two idiomorphs may include one or more genes (Turgeon and Yoder 2000). The MAT genes appear to code for transcription factors, which in turn regulate the transcription of genes pertaining to the synthesis of hormones, proteins and compounds directly relevant to sexual reproduction (Pöggeler 2000). The two MAT idiomorphs may be identified by their DNA-binding motifs: whereas MAT1-1 contains an α-box motif, MAT1-2 contains an HMG (high-motility group) box (Arie et al. 1999). This observation facilitates the discovery of mating type genes in new species, because the remainder of the sequence is often highly divergent (Taylor et al. 1999).

Within the MAT1-1 idiomorph, the α-box-containing gene is usually named MAT1-1-1 (Turgeon and Yoder 2000), and may be associated with two additional genes, called MAT1-1-2 and MAT1-1-3. Although all three genes appear to play critical roles in sexual reproduction when present (Klix et al. 2010), mutant strains of Neurospora crassa carrying MAT1-1-1 but with deletions of both MAT1-1-2 and MAT1-1-3 are capable of successful, albeit reduced reproduction (Ferreira et al. 1998). Alpha boxes are a general family of transcription regulators that are found in association with a wide variety of other proteins (Günther et al. 1998). In contrast, the MAT1-2 idiomorph consists of a high-motility group (HMG) protein, usually termed MAT1-2-1 (Turgeon and Yoder 2000) and may include a second gene, MAT1-2-2 (Pöggeler and Kuck 2000; Staben and Yanofsky 1990). The HMG box is a common motif in the
DNA-binding domain of diverse proteins, including transcription factors (Bianchi and Agresti 2005) and the mammalian sex-determining protein SRY (Dubin and Ostrer 1994).

Despite containing highly conserved amino acid motifs (Coppin et al. 1997), the overall sequences of the MAT genes are poorly conserved overall (Yun et al. 2000). Further complicating the assignment of mating type using molecular methods, both the MAT1-1-3 and MAT1-2-1 always contain an HMG box (Butler 2007; Yun et al. 2000). The high conservation of the HMG and/or the α-box portion(s) and the lack of conservation of the remaining portion of the gene may allow inadvertent assignment of the incorrect mating type to an isolate of interest. However, this error can be avoided because, to date, MAT1-1-2 and MAT1-1-3 have always been found in association with MAT1-1-1 (Lin and Heitman 2007), whereas MAT1-2-1 may be found either with an α-box (in the case of homothallic strains with MAT1-1-1) or without (such as in heterothallic strains).

Because of the poorly conserved nature of the nucleotide sequences of these mating type genes between species, the identification of MAT gene sequences without genomic information can be challenging. However, the identification of the MAT1 locus is facilitated in the Pezizomycotina because the MAT gene(s) are frequently located in between two protein-coding genes: APN2 (encoding a DNA lyase) and SLA2 (encoding a protein that plays a role in cytoskeletal assembly) (Butler 2007). However, most of the species in which this pattern has been studied belong to the Sordariomycetes, although some unpublished research from this lab (Y. Deng & T. Hsiang, personal communication) has found that this synteny is less prevalent among the Dothideomycetes; therefore, there may be class-level differences in the conservation of this organizational pattern.
Even in homothallic species that have MAT idiomorphs located in two entirely separate locations within the genome, a partial sequence of one of these usually-adjacent proteins may be located proximally to each MAT1 locus (Rydholm et al. 2007). The most common orientation is for the SLA2 sequence to be upstream of the MAT gene(s), with APN2 located immediately downstream (Butler 2007). In addition to SLA2 and APN2, there are other genes frequently associated with the MAT region in the Sordariomycetes. Two additional proteins, an anaphase-promoting complex protein (APC5), and cytochrome c oxidase (COX13), are frequently identified within the vicinity of the MAT1 locus (Butler 2007). For simplicity, the genes typically located immediately up- and down-stream of the MAT1 locus in other Sordariomycetes (SLA2, APN2, APC5, and COX13) will be collectively referred to herein as "flanking genes."

5.1.3 Sexual reproduction in M. nivale and M. majus

The sexual stage of M. nivale sensu lato has been described, but some reports of its frequency are inconsistent. Where sexual reproduction has been described, M. nivale sensu lato has been reported to form perithecia containing numerous asci 75 \( \mu \text{m} \) in length and 7-8 \( \mu \text{m} \) in width (Stevens 1918), containing eight ascospores. The ascospores are hyaline and ellipsoidal, with 0-3 septa and are 10-17 \( \mu \text{m} \) in length and 3.5-4.5 \( \mu \text{m} \) in width (Booth 1971). When the teleomorph of M. nivale sensu lato was first described as Calonectria nivalis, Schaffnit described the conidia of the corresponding asexual stage as ranging in length between 14.6 to 25. 2 \( \mu \text{m} \) in length and between 3.2 to 4.3 \( \mu \text{m} \) in width, with most conida possessing three septa (1913). The identification of C. nivalis from several cereal hosts, as well as this description of conidial size, does not clearly establish whether the anamorph was M. nivale or M. majus (Glynn et al. 2005); however, the predominance of tri-septate conidia is more similar to M. majus than to M. nivale.
Bennett (1933) observed the formation of immature perithecia only on sterilized plant fragments in the lab, as opposed to plant tissue in the field, and only when freshly-isolated fungal strains were studied. In contrast, Wollenweber reported that approximately 10% of isolates produced perithecia in the lab and that those strains that did form sexual structures produced fertile perithecia in abundance on different substrates (as cited in Bennett 1933). Gordon (1952) reported that single conidial cultures were self-fertile (i.e. that *M. nivale* sensu lato is homothallic), and that perithecia were not present in the field in Canada; however, mature perithecia were observed on wheat in North America in 1966 (Cook and Bruehl). The sexual stage of *M. nivale* sensu lato has been frequently reported in Europe (Cook and Bruehl 1966).

In more recent experiments, which distinguished between *M. nivale* and *M. majus* (then considered conspecific varieties), both species were found to produce fertile perithecia in the lab (Lees et al. 1995; Litschko and Burpee 1987; Parry et al. 1995). However, *M. majus* appears to produce perithecia more readily than *M. nivale* (Lees et al. 1995). Single-spore isolates of both species produce perithecia in the lab (Lees et al. 1995; Parry et al. 1995), and thus both species have been described as potentially homothallic. A further distinction was observed among *M. nivale* isolated from either cereal or turfgrass when isolates from cereal, but not turfgrass, produced perithecia in the lab (Smith 1983); however, fertile perithecia were obtained from paired *M. nivale* turfgrass isolates in a later experiment (Litschko and Burpee 1987). In this experiment, pairings of isolates collected from both cereal and turfgrass isolates (and cross-pairings of isolates collected from each host) yielded fertile perithecia (Litschko and Burpee 1987). This observation, in addition to the high degree of genetic variation in *M. nivale* relative to *M. majus*, has led some to suggest that *M. nivale* may also reproduce heterothallically (Lees et al. 1995; Mahuku et al. 1998). Neither the MAT nor the flanking genes from *M. nivale* and *M.
majus have been described, but the identification of these genes within this species may help to clarify some of the confusion regarding whether *M. nivale* is truly homo- or heterothallic.

### 5.1.4 Objectives

The primary objective of this project was to identify mating-type genes within multiple genomes of *M. nivale* and *M. majus* to assess whether these species are homo- or heterothallic, and whether they undergo sexual reproduction. The presence of mating type genes in the *M. nivale* and *M. majus* isolate collections was first assessed using primers developed for other fungal species, and then from whole-genome sequenced isolates of *Microdochium nivale* and *M. majus*. Mating crosses were also performed in the lab to investigate possible cross-fertility within and between species.

### 5.2 Materials and Methods

#### 5.2.1 Test of mating type primers based on conserved sequences

Published MAT1-1-1 and MAT1-2-1 primers (Table 5.1) were tested with DNA from nine isolates of *M. nivale* and two isolates of *M. majus* (Table 5.2). Each reaction was repeated twice, using annealing temperatures of 55 and 53 °C, and using a final concentration of 0.4 mM of Mg$^{2+}$ in each reaction mixture. The thermocycling program was as described for the ITS amplification (Section 2.2.4).

Mating-type primers for MAT1-1 and MAT1-2 were designed by manually selecting conserved regions in the alignment of nucleotide sequences collected from GenBank (Table 5.3), and aligning and visualizing the sequences as described (Section 2.2.6). The primers that were designed in this manner (Table 5.4) were tested using the isolates summarized in Table 5.1. All primers were tested using the PCR protocol and thermocycling conditions described for the ITS amplification.
amplification (Section 2.2.4), with the exception that annealing temperatures between 50-62 °C were tested for the primer sets. The amplicons from seven isolates amplified with primers designed in this manner were sequenced and analysed in the manner described in Chapter 2 (Section 2.2.4).

5.2.2 Identification of putative MAT1 loci and flanking genes and screening of isolate collection

The amino acid sequences of MAT1-1-1, MAT1-1-2 MAT1-1-3, MAT1-2-1, and the flanking genes SLA2 and APN2 from several species of filamentous ascomycetes (Table 5.2) were used to query both the assembled scaffolds and the predicted gene sequences from the genomes of M. majus isolate 99049 and M. nivale isolates 11037 and 12262 (Chapter 3). Primers were designed for the putative sequences of these genes by selecting regions amenable to primer design (Section 2.2.6).

Four primers, Mn_MAT2_3347F, Mn_MAT2_3871R, Mn_lyase_838F, and Mn_SLA2_23R were designed based on an early assembly of the M. majus 99049 genome, which was the first Microdochium sp. genome available (Chapter 3). The MAT1-2-1 primers Mn_MAT2_3347F and Mn_MAT2_3871R were tested using DNA from both M. nivale and M. majus (including the sequenced isolate) using the PCR conditions described for the ITS amplification (Section 2.2.4). The Mn_lyase_838F and Mn_SLA2_23R primer set was first tested using DNA from the sequenced isolate and using the PCR mixture described for the RPB2 amplification (Section 2.2.4), with a thermocycling protocol consisting of an initial 5 minute denaturation at 95 °C followed by 35 cycles of a 1 minute denaturation at 95 ºC, a 1 minute annealing period at 55 °C, and a 3 minute extension at 72 ºC. The final extension consisted of a
10 minute period at 72 °C. This protocol was modified by increasing the annealing period to 2 minutes and the extension period in 1-minute increments up to 6 minutes in length.

This reaction was also attempted using the Expand LT (Expand Long Template PCR system; Roche, Laval, Quebec) enzyme mixture. The reaction was performed using buffer mixture #2 (containing 27.5 mM of Mg²⁺) and according to the manufacturer's directions with the exception that a final concentration of 500 nM, rather than 300 nM, of each primer was included. The basic thermal cycling procedure consisted of an initial denaturation of 2 minutes at 94 °C, followed by 10 cycles of a 10 s denaturation at 94 °C, a 30 s annealing period at 50 °C, and an 8 minute extension period at 68 °C. This was followed by 20 cycles with the same parameters, except that the duration of the extension period was increased by 20 s each cycle. These cycles were followed by a final extension at 68 °C for 7 minutes. This cycle was modified by varying the annealing temperature up to 63 °C and by increasing the extension temperature to 70 °C to reduce non-specific amplification.

Later, based on a more complete assembly of the *M. majus* 99049 genome (Chapter 3), the primer Mn_SLA2_1156F was designed. This primer was paired with the Mn_lyase_838F primer using the Expand LT amplification conditions described above. In addition to these reactions, the Mn_SLA2_1156F primer was also paired with the Mn_MAT2_3871R primer and the Mn_MAT2_3347F primer was paired with the lyase_838R primer to confirm the synteny of this region. These reactions were performed using the Expand LT conditions described above.

The alignments for each set of primers designed are found in Appendices 5.1 through 5.4, and a list of the primers ordered and tested is found in Table 5.4. All *M. majus* isolates available in the local lab collection were screened with the primer set Mn_MAT2_3347F and Mn_MAT2_3871R to determine whether the putative MAT1-2-1 gene identified in *M. majus*
isolate 99049 could be found in all isolates. The region spanning between the SLA2 to APN2 genes was also tested in all available isolates.

When the *M. nivale* 11037 genome became available, the primer set Mn_MAT2_20F and Mn_MAT2_727R was designed to amplify a partial sequence of the MAT1-2-1 gene. These primers were tested using the PCR conditions described for the amplification of ITS (Section 2.2.4). The primer set Mn_SLA2_1156F and Mn_APN2_700R was designed to amplify a region spanning from the end of the putative SLA2 gene to the beginning of the putative APN2. In addition, the primers Mn_SLA2_1156F and Mn_MAT2_727R, and separately, Mn_MAT2_20F and Mn_APN2_700R were paired to test the syntenic of the region surrounding the putative MAT1 locus. All reactions involving either the APN2 or the SLA2 primers were performed using the Expand LT enzyme as described above. To determine whether the putative MAT1-2-1 gene identified in *M. nivale* isolate 11037 could be amplified in all *M. nivale* isolates, a variety of *M. nivale* isolates from different geographic and host origins were screened with the primer set Mn_MAT2_20F and Mn_MAT2_727R (Table 5.8).

The MAT sequence data from *M. nivale* isolate 12262 was combined with the data from the previously-sequenced *M. nivale* 11037 and *M. majus* 99049, and was used to design the primer sets Mic_MAT2_198F and Mic_MAT2_676R and Mic_SLA2_92F and Mic_APN2_32R that were designed to amplify the putative MAT1-2-1 and the region between SLA2 and APN2. The amplification protocols for these primer sets were as described for the equivalent reactions performed using the primers designed with the data from *M. nivale* isolate 11037. These primer sets were used to amplify the region spanning from the end of the putative SLA2 through to the beginning of APN2, and to amplify the putative MAT1-2-1 region in DNA isolated from both *M. nivale* and *M. majus*.
5.2.3 Mating experiments

Wheat straw was collected from the dairy barn at the University of Guelph. The straw was cut into segments 2-3 cm long, and these fragments were autoclaved three times at 121 °C for 20 minutes. Following sterilization, the wheat straw was placed into the centre of a Petri plate containing either PDA (experiment #1) or water agar (experiment #2). Water agar was prepared by combining 20 g of Bacto agar (Becton, Dickinson and Company, MD, USA) with 1 L of water and autoclaving for 20 minutes at 121 °C. The straw was inoculated with two plugs of agar cut from the actively-growing margin of a colony. A list of the isolates used and the crosses performed in these experiments is found in Table 5.9. The agar plugs were placed at opposite ends of the straw, and the plates were sealed with Parafilm and incubated in the dark in loosely-sealed plastic storage bins (KIS Omni Box, 9L, Milton, ON). In experiment #1, the plates were stored at 10 °C, and in experiment #2, replicate plates were incubated at 5, 10, 15, and 20 °C. Plates were checked weekly for the presence of perithecia.

Where perithecia were present, a single perithecium was gently removed from the surface of the wheat straw and was placed in 100 µL of sterile dH2O. The perithecium was vortexed for 15 s, and then allowed to settle at room temperature for 15 minutes. An aliquot of 50 µL was then viewed under 400 x magnification to check for the presence of ascospores. Ascospores of Microdochium spp. are shorter and thinner than conidia (Booth 1971). The average size and number of septa from ten ascospores was recorded for each of the isolates producing ascospores. When putative ascospores were observed, the spore suspension was streaked on a PDA plate and incubated at room temperature for 24 h. After 24 h, single germinated ascospores were identified at 100 x magnification and were transferred to fresh PDA plates. Mycelium was harvested from ten single-ascospore cultures in the manner described in Chapter 2.
To ensure that ascospores were obtained and not conidia, their size was compared to those of conidia. The average dimensions (length, width, and number of septa) were recorded for ten conidia each for eight mono-conidial colonies of *M. majus*. Conidiation was induced by incubating inoculated PDA plates overlaid with cellophane under constant light for up to one week at room temperature. Conidia were collected by pipeting 2 mL of sterile dH₂O onto the surface of the plate, scraping gently with a sterile glass rod, and viewing the resulting conidial suspension under 400x magnification.

To investigate the similarity of the genotypes of the single-ascospore (SA) cultures to the "parent" culture and to the other "sibling" cultures, the DNA isolated as described above was assessed in two ways. First, the presence or absence of the putative MAT1-2-1 gene identified as described above was assessed using the primers Mic_MAT2_198F and Mic_MAT2_676R. The genetic diversity was also assessed using the ISSR and SSR primers listed in Table 4.3 using the PCR protocol described in Chapter 4. The banding patterns obtained for the ten SA cultures and the parent culture were also compared to those for an additional seven *M. majus* isolates. The resulting binary matrix was used to construct a bootstrapped UPGMA tree (Section 4.2.4).

### 5.2.4 Comparison of *Microdochium* sp. with other species

The synteny of the MAT region in several fungal species were compared to that observed in the genomes of *M. majus* isolates 99049 and 10095, *M. nivale* isolates 11037, 12262, and 10106, and *M. bolleyi* isolate 07020. The genomes used in these comparisons were all members of Xylariales including *Daldinia eschscholtzii* (downloaded from JGI), *Annulohypoxylon stygium* (courtesy of B. Xie, Fujian Agriculture and Forestry University), two species of *Hypoxylon*, *Pestalotiopsis neglecta* (courtesy of K. Watanabe, Tamagawa University), and *Pestalotiopsis* sp.
The putative genes for APN2, SL2A, APC5, COX13, and, where present, MAT1-1-1, MAT1-1-2, MAT1-1-3, and MAT1-2-1 were identified by performing BLASTx searches using published protein sequences for these genes (Table 5.2) against BLAST databases of the assembled sequences of each genome. A map of the MAT region was constructed for each of the genomes studied. For the Microdochium spp., the putative gene sequences identified were also searched against the predicted gene sets for each species.

For those genomes where putative MAT genes were found in locations distant to the typical flanking genes, the regions 10,000 bp both up- and down-stream of the putative mating type genes were examined to determine whether there were any other putative genes that were common to these loci. The genes found within these regions were tentatively identified in the other genomes tested to search for commonalities in their location relative to the other genes of interest.

### 5.3 Results

#### 5.3.1 Test of published mating type primers and redesigned universal primers

Primers designed to amplify the mating type genes MAT1-1-1 and MAT1-2-1 from several different species of Ascomycetes were tested on 11 isolates of *M. nivale* and *M. majus* to identify the presence of these genes in these species. Using the published MAT1-2-1 primers ChHMG1 and ChHMG2, a faint band of the predicted size (approximately 300 bp) was produced from three *M. nivale* isolates. The MAT1-1-1 Falpha1 and Falpha2 primer set also amplified several weak bands (approximately 2,000, 1,000, and 300 bp in length) from both the *M. nivale* and the *M. majus* isolates. Despite optimization attempts, a single band was not obtained from either of these primer sets.
With the goal of producing a single band in sufficient quantities for sequencing, new MAT1-1-1 and MAT1-2-1 primers were designed from conserved regions of MAT sequences downloaded from GenBank (Table 5.3). Despite testing all of the primers described in all possible combinations (i.e. all forward primers for a single gene were tested with all of the reverse primers available), putative amplicons of neither MAT1-1-1 nor MAT1-2-1 could be amplified from either species of interest using previously-published data.

5.3.2 Identification of putative mating type and flanking genes

The sequencing data of *M. majus* isolate 99049 provided the first *Microdochium* genome obtained (May 2011). When the raw data from this genome were assembled (Chapter 3), an attempt was made to identify putative sequences for the MAT1-1-1 or the MAT1-2-1 genes using sequences collected from GenBank (Table 5.6). In the early assembly of the *M. majus* genome available at that time, putative matches for MAT1-2-1, SLA2, and APN2 were identified on separate contigs. A putative match for MAT1-1-1 was not found. More details regarding the identification of the flanking genes in this isolate are found in Appendix 5.11.

A set of primers, Mn_MAT2_3347F and Mn_MAT2_3871R was designed to amplify the putative MAT1-2-1 sequence. The MAT1-2-1 primers designed with the Mm99049 genome successfully amplified a single band of the predicted size with all of the *M. majus* isolates tested, but failed to yield an amplicon from any of the *M. nivale* isolates tested. When the amplicon of *M. majus* isolate 10148 was sequenced in the forward direction, the resulting sequence was 100% identical to the putative MAT1-2-1 sequence from the genome data across the 441 bp sequence obtained. None of the *M. nivale* isolates tested could be amplified using these primers.
A second set of MAT1-2-1 primers, Mn_MAT2_3356F and Mn_MAT2_3979R were prepared based on the putative MAT1-2-1 sequence identified in the genome. All of the *M. majus* isolates tested with this primer set produced a single band of the predicted size (623 bp), but, again, none of the *M. nivale* isolates tested showed a band with this primer pair.

The flanking genes SLA2 and APN2 were tentatively identified in the *M. majus* genome by querying sequences collected from GenBank (Table 5.6) against the early assembly of the *M. majus* genome. These genes were identified based on their high sequence identity with the query sequences (Table 5.7). However, in this early assembly, each of these genes and the putative MAT1-2-1 gene were found on separate contigs, and as a result the synteny of this region that was observed in other Sordariomycetes (Figure 5.1) could not be confirmed using the genomic data alone.

When an improved assembly of the *M. majus* genome became available (December 2011), the putative SLA2, APN2, and MAT1-2-1 genes previously identified were found to be contiguous. However, these genes were arranged in a different configuration from that observed for many other Sordariomycetes (Figure 5.1). This new information was used to design a new primer, Mn_SLA2_1156F which was paired with the Mn_APN1_838F primer which, serendipitously, functioned as a reverse primer as the putative lyase gene was downstream of SLA2 (rather than upstream, as predicted from other species), and possessed an inverted orientation relative to other species. This combination of primers yielded a single band of the predicted size (8 kb) for all of the *M. majus* isolates with which it was tested. The orientation of the MAT1 region and the flanking genes as proposed based on the whole-genome assembly of *M. majus* isolate 99049 was confirmed by amplifying the region from the tail end of the SLA2 gene to the MAT1-2-1 gene (using primers Mn_SLA2_1156F and Mn_MAT2_3871R) and the
region from the MAT1-2-1 gene to the beginning of the APN2 gene (using primers Mn_MAT2_3347F and lyase_838F). Both reactions yielded single bands of the predicted sizes (4.5 kb and 2.3 kb) for all of the \textit{M. majus} isolates tested, regardless of their geographic origin.

The combination of primers described above that successfully amplified the APN2 through the SLA2 cassette as well as the MAT1-2-1 gene and intergenic regions were applied to DNA from \textit{M. nivale} isolates from grass and wheat, and from both Europe and Canada. However, these primers failed to amplify the genes of interest under any of the conditions attempted. All of the putative MAT and flanking genes identified in \textit{M. majus} were also found among the genes predicted using AUGUSTUS (Chapter 3). A summary of the similarity of these predicted genes to the best match among the sequences used to query against these sequences is found in Table 5.7.

When the genome of \textit{M. nivale} isolate 11037 became available (February 2012), the putative MAT1-2-1, APN2, and SLA2 genes were identified in both the genome assembly and in the predicted gene set by using both the sequences listed in Table 5.6 and the putative MAT1 locus and flanking genes identified in \textit{M. majus} isolate 99049. A summary of the similarity of these predicted genes to the best match among the sequences used to query against these sequences is found in Table 5.7. A description of the similarity of these sequences to the putative orthologs in \textit{M. majus} is described in Section 5.2.3. As in the later assembly of \textit{M. majus}, all three genes were found on a single contig, and there was no match for MAT1-1-1. The configuration of the MAT1 locus and flanking genes in \textit{M. nivale} 11037 was identical to that predicted for \textit{M. majus} 99049, with the exception of minor variations in sequence length (Figure 5.2).
Primers were designed for the APN2 (Mn_APN2_32R), SLA2 (Mn_SLA2_357F), and MAT1-2-1 (Mn_MAT2_20F and Mn_MAT2_727R) genes identified in the genome of *M. nivale* isolate 11037. Following optimization, the SLA2 and APN2 primer set was used to successfully amplify a band of the predicted size (7 kb) in all of the *M. nivale* and *M. majus* isolates tested. The MAT1-2-1 primers yielded a band of the predicted size (707 bp) from the sequenced *M. nivale* isolate 11037, but produced inconsistent results with many of the other *M. nivale* isolates tested; only 59% of the 92 isolates tested were amplified. There was no apparent relationship between an isolate's host plant or general geographic origin and its amplification by the MAT1-2-1 primers. A summary of the number of isolates tested and the number of these which could be amplified using these primers may be found in Table 5.7. None of the *M. majus* isolates could be amplified with the *M. nivale* MAT1-2-1 primers.

The genome of *M. nivale* isolate 12262 was chosen for whole-genome sequencing (Chapter 3) because this isolate was among the group of *M. nivale* isolates that could not be amplified using the primers designed from *M. nivale* 11037. The putative MAT1-2-1, APN2, and SLA2 genes were identified in the draft genome assembly and in the predicted gene set by using both the sequences listed in Table 5.6 and the putative MAT1 locus and flanking genes identified in *M. majus* isolate 99049 and in *M. nivale* 11037. A summary of the similarity of these predicted genes to the best match among the sequences used to query against these sequences is found in Table 5.7.

When the genome of *M. nivale* isolate 11037 became available (February 2012), the putative MAT1-2-1, APN2, and SLA2 genes were identified in both the genome assembly and in the predicted gene set by using both the sequences listed in Table 5.6 and the putative MAT1 locus and flanking genes identified in *M. majus* isolate 99049. A description of the similarity of
these sequences to the putative homologs or orthologs in the other *Microdochium* genomes described is described in Section 5.2.3. As described for the other *Microdochium* genomes, all three genes were found on a single contig and shared the same synteny (Figure 5.2) with only minor variation in sequence length. No match for MAT1-1-1 was observed.

The final sets of MAT1 locus and flanking gene primers were designed by combining the data from all three of the genome sequences available. The primers were Mic_SLA2_92F and Mic_APN2_32R (spanning the region between the 3' end of SLA2 to the 5' end of APN2), and Mic_MAT2_198F and Mic_MAT2_676R, which fell within the putative MAT1-2-1 sequence.

The "universal" MAT1-2-1 primers successfully amplified the putative MAT1-2-1 gene from all of the *M. nivale* and *M. majus* isolates with which they were tested, including both those which could and could not be amplified with the Mn_MAT2_20F and Mn_MAT2_727R primer set (designed using the March 2012 genome assembly of *M. nivale* isolate 11037). The SLA2 through APN2 region was also successfully amplified in all *M. nivale* and *M. majus* isolates tested using this "universal" primer set.

### 5.3.3 Mating experiments

In the first mating experiment, isolates of *M. majus* and *M. nivale* (Table 5.9) were inoculated on sterilized wheat straw placed on PDA and incubated in the dark at 10 °C for 6 months. All possible crosses were performed with the isolates included, including self by self. After 16 weeks, perithecia were visible on all of the plates inoculated with at least one isolate of *M. majus*. In crosses including one isolate of *M. majus* and one isolate of *M. nivale*, perithecia were found only on half of the straw closest to the *M. majus* inoculum. In crosses including two different isolates of *M. majus*, or in plates including a self by self *M. majus* cross, perithecia were
found throughout the plant tissue. The perithecia were checked for asci and ascospores for an additional 10 weeks following the initial observation of perithecia, but for the remainder of the experiment, neither asci nor ascospores were observed.

In the second experiment, water agar, rather than PDA, was used as the medium in the Petri dishes containing sterilized wheat straw. The isolates included in these experiments are summarized in Table 5.9. The isolates were plated on the wheat-straw plates in all possible crosses, including self by self. Four copies of each possible cross were prepared, and one copy of each plate was stored at 5, 10, 15, or 20 °C in the dark.

Table 5.10 summarizes the results of the mating crosses after two months of incubation. In crosses containing two isolates of *M. majus* (including self by self crosses), and incubated at 20 °C, perithecia (Figure 5.4) were observed in 80% of the crosses, including in all four of the self by self crosses. Among the eight cases in which perithecia were produced, asci with ascospores were observed in six out of the eight crosses (Table 5.10). In plates incubated at 20 °C containing crosses of *M. majus* with *M. nivale*, perithecia were observed on twelve of the twenty plates prepared, and of the twelve plates with perithecia, asci and ascospores were observed in six cases. All four of the *M. majus* isolates produced perithecia and ascospores in at least one pairing, and when perithecia were produced, they were clearly located on the side of the wheat straw that was closest to the *M. majus* inoculum (Figure 5.4). Among the plates inoculated with *M. nivale* only, perithecia were produced on only two plates, and neither asci nor ascospores were observed.

The production of perithecia was observed less frequently at 15°C. At this temperature, perithecia were observed in only six of the ten plates inoculated with two isolates of *M. majus*, and of these six, ascospores were observed in only three. In the plates containing *M. majus*
crossed with *M. nivale*, only eight of the twenty plates contained perithecia, but of the plates with perithecia, ascospores were observed in eight. None of the plates inoculated with two isolates of *M. nivale* yielded perithecia.

At 10 °C, only three of the plates inoculated with two isolates of *M. majus* yielded perithecia, and ascospores were observed in only one case. None of the *M. majus* by *M. nivale* plates yielded perithecia, but three of the plates inoculated with two isolates of *M. nivale* yielded perithecia, although ascospores were not observed in these structures. Neither ascospores nor perithecia were observed on any of the plates incubated at 5°C.

To confirm that the spores observed in the perithecial tissue were ascospores as opposed to conidia, their sizes were recorded. A total of ten ascospores for three of the self by self *M. majus* crosses were measured (Appendix 5.9). The putative ascospores were (12.0-)14.4 to 19.2 (-24.0) µm in length and (2.4-) 3.6 to 4.8(-6.0) µm in width, with 1-3 (predominantly 3) septa observed. The average size of the putative ascospores did not differ between different isolates. The average conidial dimensions were also recorded for a total of eight mono-conidial colonies of *M. majus* to compare the average appearance of *M. majus* conidia to the putative ascospores. For the eight isolates tested, the average dimensions of the conidia were 26 by 5 µm, with between 0-3 septa.

The DNA from ten single-ascospore (SA) cultures derived from the apparently homothallic *M. majus* isolate 99049, as well as DNA from this "parent" culture, was isolated and analysed using the Mic_MAT2_198F and Mic_MAT2_676R primer set to determine whether these spores contained the putative MAT1-2-1 gene identified in the other *M. nivale* and *M. majus* isolates examined. The DNA from all ten of the SA cultures, as well as the DNA from the parent culture, produced a single band of the predicted size (478 bp) in all of the samples,
suggesting that the putative MAT1-2-1 gene was indeed present. This DNA, in addition to DNA isolated from an additional seven *M. majus* isolates (Table 5.5), was also analysed using the ISSR and SSR primers listed in Table 4.3. All seven of the other isolates were also successfully amplified by the Mic_MAT2_198F and Mic_MAT2_676R primer set. This subset of isolates were chosen because 99061 was collected at the same time and location as isolate 99049 (in Atwood, ON, in 1999); isolates 12043-12046 were all collected from the same field (in Ottawa, ON, in 2012); and isolates 10098 and 10099 were from Europe and thus intended as out-groups.

For the ISSR primers BHY(AGC)₅ and DD(CCA)₅ and for the SSR primer sets MnSSR_CT₅, MnSSR_GAT₆, MnSSR_916061, and MnSSR_924078, identical banding patterns were observed with the DNA isolated from all 10 of the SA cultures as well as from the "parent" culture. For the ISSR primer (CAC)₅, the DNA from eight of the ten SA cultures and the parent culture yielded a single band of approximately 700 bp. No band was observed for the DNA from the other two SA cultures. For the SSR primer MnSSR_GC₅, five of the SA cultures yielded identical banding patterns with a total of four bands, while four of the SA cultures and the parent culture yielded a three-band pattern that was missing a band of approximately 1.6 kb that had been found in the four-band isolates. The remaining SA culture was not amplified by these primers. The relationships between the SA cultures, the parent culture, and the seven other isolates were visualized by constructing a UPGMA tree with 100 bootstrap replications (Figure 5.6). In this tree, all SA cultures and the parent culture grouped into a single clade with a bootstrap value of 82%. The isolate collected at the same time as the parent culture (99061), did not group with these isolates. The four Ottawa isolates were grouped into a single clade with 37% bootstrap support that also included European isolate 10098.
5.3.4 Comparison of the *Microdochium* MAT1 locus to that of other species

The synteny of the MAT1 locus and the flanking genes was assessed using whole-genome data from *Microdochium nivale, M. majus, M. bolleyi*, and seven other species within the Xylariales (Table 5.11). In addition to the flanking genes SLA2 and APN2 described above, the genes APC5 and COX13, also associated with the MAT1 locus in many species (Butler 2007), were also identified. The sequences used to query against these genomes are summarized in Table 5.6.

Putative sequences for all four of the flanking genes and MAT1-2-1 were identified in all 13 genomes examined. Unlike the *Microdochium* genomes, with the exception of *Pestalotiopsis* sp., the putative MAT1-2-1 sequences identified in the other Xylariales genomes were not located in proximity to the putative flanking genes. Only very weak matches for MAT1-1-1 and MAT1-1-2 (e-values of 0.02 or greater) were identified in any of the Xylariales genomes. The MAT1-1-3 query sequence, which contains an HMG box, had at least two matches with an e-value less than 1e^{-05} in all of the species examined, including the *Microdochium* genomes. The same matches were identified for the MAT1-2-1 query sequence. For all 12 genomes, all of the sequences identified as putative MAT1-2-1 and / or MAT1-1-3 genes were queried against the GenBank database (BLASTx) (Table 5.11). For all twelve genomes, the best matches for these sequences were putative HMG-box containing proteins. For those species where the putative match to MAT1-1-3 / MAT1-2-1 was distant from any of the flanking genes, it was not possible to determine whether the sequence identified was likely one of these two mating-type genes or if it was more likely to be a different HMG-box containing protein. In the *Microdochium* genomes, as well as in *P. theae*, the HMG-box containing protein found between the flanking genes was
tentatively assigned the identity of a putative mating-type gene. The apparent absence of MAT1-1-1 led to the classification of this mating type gene as MAT1-2-1.

All of the sequences identified as matches to the four flanking genes were also further investigated (Table 5.11) to determine whether there were any putative duplicated sequences for the four flanking genes. Although there was more than one match for each of the flanking genes in many of the genomes studied, in most cases these additional sequences appeared to be unrelated to the flanking genes. In both *M. majus* 99049 and *M. majus* 10095, secondary matches (in addition to those found in proximity to the putative MAT region) for cytochrome c oxidase were identified on a separate scaffold; however, in both cases, this sequence was not found in the proximity of any of the other secondary matches to the other MAT-related genes identified. A similar trend was observed for *M. nivale* isolate 12262, where a secondary match for APC5 was identified on a separate scaffold relative to the other sequences of interest.

The relative genomic organization of the proposed mating region, including the flanking sequences, for all of the genomes studied is depicted in Figure 5.2. The orientation of the flanking genes relative to the putative MAT gene was conserved in all of the *Microdochium* spp. studied, but differed in all seven of the other genomes studied, including the three *Hypoxylon* sp. genomes.

### 5.4 Discussion

In this Chapter, the putative mating type regions of *Microdochium nivale, M. majus,* and *M. bolleyi* were identified from the whole-genome data from these three species based on comparison to other published sequences. Primers were designed from the sequences identified and were used to screen the isolate collections of *M. nivale* and *M. majus,* and mating crosses
were performed in an attempt to further explore the conditions required for sexual reproduction to occur.

Based on earlier reports that *M. majus* and *M. nivale* are homothallic, both MAT1-1-1 and MAT1-2-1 were predicted to exist within the genomes of all isolates of these species. The initial search for the MAT genes was conducted using the published primer sets Falpha and ChHMG in an attempt to amplify MAT1-1-1 and MAT1-2-1, respectively, from the DNA of *M. nivale* and *M. majus*. However, inconsistent results were obtained with these primers, and a single band of the predicted size was not produced by either primer set. The lack of amplification with these primers could be explained by considering that MAT genes are highly variable at the nucleotide level, and that *Microdochium* is a member of the Xylariales, whereas the MAT1-1-1 primers were designed for use in *Fusarium* sp. (a member of the Hypocreales), and the MAT1-2-1 primers were designed for use in *Colletotrichum higginsianum* (Chen et al. 2002), a member of the Glomerellales.

To address these issues, several sets of "conserved" MAT1-1-1 and MAT1-2-1 primers were then designed based on MAT1-1-1 and MAT1-2-1 sequences collected from GenBank. The goal was to design new primers that were biased towards species that may be more closely related to *Microdochium* than those that were used to design the previously-tested primers. Although some of the primers did amplify bands of varying sizes from *M. nivale* and *M. majus*, when these bands were sequenced, they displayed low homology with any of the available sequences in the GenBank database and shared no apparent relationship with the MAT sequences that had been used to design the primers. These results suggested that these primers were probably annealing to sequences in the genome that are unrelated to the sequences of
interest. Based on these inconsistent and disappointing results, the search for MAT genes was postponed until the whole-genome data became available (Chapter 3).

When the genome sequencing data of *M. majus* isolate 99049 were first assembled (May 2011), a putative MAT1-2-1 sequence was identified on a single contig based on its homology to the query sequences used. Surprisingly, no putative match to MAT1-1-1 was identified; however, especially with this single and preliminary genome sequence, it was not possible to determine whether the MAT1-1-1 gene was truly missing, whether it was merely situated in a poorly-assembled region of the genome, or whether it bore too little homology to the query sequences to be identified in this manner.

Despite this confusion, primers were designed to amplify this putative MAT1-2-1 sequence. The primer set produced a single band of the predicted size from all of the *M. majus* isolates tested, but did not yield consistent results with DNA from *M. nivale*. Sequencing the PCR product from one of the *M. majus* isolates produced a sequence that was identical to the putative *M. majus* MAT1-2-1 sequence identified from the genome. Although several different sets of conditions were tested, the *M. majus* MAT1-2-1 primers did not produce bands of the predicted size from DNA of *M. nivale* isolates. This inability to amplify the gene of interest was likely due to sequence differences between *M. nivale* and *M. majus*, and the genomic sequences showed that there were polymorphisms in the priming region.

In an attempt to investigate the MAT region of *M. nivale* before its genome sequence became available, a strategy employed by other researchers (Putman et al. 2011) was employed. In this method, the genes APN2 and SLA2, predicted to be located immediately up- and downstream, respectively, from the putative MAT1-2-1 sequence (Figure 5.1), were identified in the *M. majus* genome. The presence of these genes within the vicinity of the putative MAT1-2-1
sequence also affirmed that this was the MAT locus in *M. majus*. Although sequences sharing high homology to both sequences of interest were identified, all three sequences of interest (the two flanking genes and the putative MAT1-2-1) were located on different contigs in the May 2011 assembly of *M. majus* isolate 99049, and thus the synteny of these genes relative to that in other species could not be assessed from the genomic data alone. Regardless, based on the gene order in other species (Figure 5.1), one (forward) primer was designed near the 3’ end of the APN2 sequence, and a second (reverse) near the 5’ end of the SLA2 sequence, with the goal of amplifying the region spanning from the tail end of the first flanking gene through to the beginning of the second, which was hypothesized to encompass the putative MAT1-2-1 sequence. Because the sequences of the flanking genes are more strongly conserved than those of the MAT genes (Putman et al. 2011), this method was also designed to provide access to the *M. nivale* MAT1 region. Based on other species, this region was expected to be between 4 to 10 kb in length.

Unexpectedly, the attempted amplification of the full region between APN2 through to SLA2 was unsuccessful with *M. majus* DNA, including that of the sequenced isolate 99049, even when the Expand LT enzyme system, designed to amplify fragments up to 20 kb, was used. Amplification from the 5’ end of the putative MAT1-2-1 to the 5’ end of SLA2, and from the 3’ end of APN2 through the 3’ end of MAT1-2-1, was also attempted based on the hypothesis that the full APN2 through SLA2 sequence may be longer than predicted. It was expected that at least one of these two pairings might provide a sequence short enough for successful amplification, but no product was ultimately obtained from either pairing.

When the genome of *M. majus* isolate 99049 was re-assembled (December 2011, Chapter 3), the previously-identified putative sequences of APN2, MAT1-2-1, and SLA2 were found on a
single scaffold. However, unlike in other species, the order of these genes was different from the prediction: SLA2, rather than APN2, was found to be upstream of the putative MAT1-2-1 gene and, relative to the original contig studied, the APN2 sequence was reversed. To confirm this unexpected observation, a new primer was designed at the 3’ end of SLA2, with the intention of amplifying the region between SLA2 through to APN2. As the putative APN2 sequence was reversed relative to the original contig, the previously-designed APN2 primer was used as a reverse primer. These new primers successfully amplified an 8 kb fragment from all of the *M. majus* isolates available in the local lab collection, but still failed to amplify DNA from *M. nivale*.

However, the successful PCR result confirmed the unexpected relative positions of the flanking genes and the putative MAT1-2-1 gene in *M. majus*. The overall size of this region is in agreement with the size of this region in other species. No other genes were detected in the intergenic regions between the putative MAT1-2-1 gene and the flanking genes. When predicted gene sets were prepared using the whole-genome data, both of the flanking genes and the putative MAT1-2-1 sequences identified previously were found within the predicted genes, further confirming that these sequences were indeed the desired genes of interest. In addition, as predicted from the scaffold sequences, no predicted genes were found between the flanking genes and the putative MAT1-2-1. None of the predicted genes displayed homology to MAT1-1-1.

When the whole-genome data for *M. nivale* isolate 11037 became available (February 2012), the putative SLA2, APN2, and MAT1-2-1 genes were identified on a single contig. These three genes shared the same synteny as in *M. majus* 99049. The predicted protein set for this isolate also contained all three of the sequences identified in the scaffolds, and did not contain
any predicted genes between either flanking gene and the putative MAT1-2-1. When primers were designed using the *M. nivale* 11037 putative MAT1-2-1 sequence as a template, only 59% of the *M. nivale* isolates screened yielded the predicted results. There was no apparent pattern among which isolates could or could not be amplified with respect to the isolate's host plant or continent of origin. However, those isolates that could be amplified with the MAT1-2-1 primers could also be amplified with the flanking primers (and vice-versa); similarly, the failure to amplify the MAT1-2-1 gene predicted the inability to amplify the flanking genes. This relationship led to the development of the hypothesis that those *M. nivale* isolates that could not be amplified may possess the MAT1-1-1 gene, rather than MAT1-2-1, and that the non-amplification of the flanking genes spanning the MAT1 cassette may be indicative of a size difference between this region in the amplifying and non-amplifying isolates. This hypothesis also agreed with the hypothesis that at least some isolates of *M. nivale* may be heterothallic (Lees et al. 1995; Mahuku et al. 1998). To investigate this hypothesis, the *M. nivale* isolate 12262 (which could not be amplified by any of the existing MAT-region primers) was submitted for genome sequencing.

When the genome sequence of *M. nivale* isolate 12262 was obtained, putative sequences for SLA2, MAT1-2-1, and APN2 were identified on both a single scaffold and in the predicted gene sets. In contrast to the hypothesis, MAT1-1-1 was not identified in the genome of this isolate. The sequences of these genes were aligned for all three of the sequenced *Microdochium* isolates (Appendix 5.3, Appendix 5.4, Appendix 5.5), and the priming regions for all of the primers designed and ordered to date were investigated. For all three genes, the sequence obtained from *M. nivale* isolate 12262 was more similar to that from *M. majus* 99049 than to *M. nivale* 11037. At least one mismatch for the other species was detected within the priming region.
of all of the primers ordered previously. For the two *M. nivale* isolates, the SLA2 primer had only a single mismatch at a position 5 base pairs from the 3’ end of the primer; when the scaffold sequences were compared, the full length of the region from the 5’ end of SLA2 to the 3’ end of APN2 (which includes the putative MAT1-2-1) differed by only 30 bp between these two isolates.

The inability of the flanking genes to amplify this region in *M. nivale* isolate 12262 (and other isolates), is perhaps better-explained by polymorphism in the 3’ end of the priming site. The full length of the region between the SLA2 and APN2 primers was 7,420 bp based on the scaffold sequence; if the activity of the enzyme were impaired by a poorly-available priming site caused by this mismatch in the primer, lack of amplification of the full, desired sequence would result. Primers designed based on the alignment of all three of the primer sequences (Mic_MAT2_198F and Mic_MAT2_676R for the MAT1-2-1 gene and Mic_SLA2_92F and Mic_APN2_32R for the flanking genes) successfully amplified DNA from all of the *M. nivale* and *M. majus* isolates tested. Based on this result, all of the *M. nivale* and *M. majus* isolates in the local lab collection appear to possess the MAT1-2-1 gene. The dissimilarity of the MAT1-2-1 genes identified in *M. nivale* isolates 12262 and (later) 10106, compared to that from isolate 11037 suggests that there may be two distinct alleles of the MAT1-2-1 gene among this species. Ultimately, obtaining information about the MAT genes in *M. nivale* and *M. majus* by sequencing the genomes of these species was both technically simpler, less time-consuming and less expensive than trying to obtain this information using published primers from other genera.

The discovery of MAT1-2-1 as apparently the only mating-type gene in *M. nivale* and *M. majus* was unexpected based on the reported ability of these species to undergo homothallic reproduction. The apparent absence of MAT1-1-1 from the genomes of the sequenced isolates
seems to directly contradict this claim since conventional knowledge holds that both mating types are required for ascospore production (Butler 2007), although some exceptions to this trend are known (e.g. (Menat et al. 2012)). The amplification of MAT1-2-1 from all of the *M. nivale* isolates tested, including those from both wheat and turf and from both Europe and North America suggests that, if these isolates of this species are heterothallic, they all possess the same mating type. One explanation for this imbalance could be the rarity of MAT1-1-1 in the field. As only 12 isolates of *M. nivale* from Europe were available for analysis, it is possible that, due to simple chance, none of these isolates possessed MAT1-1-1. If this mating type truly is uncommon in Europe, this imbalance could also have resulted in the complete absence of this mating type in North America if the founding population of this species was small. This lack of MAT1-1-1 in North America could also explain why perithecia have not been observed on grasses (Lees et al. 1995). Although the rarity of one mating type among populations of heterothallic species has been reported (e.g. (Linde et al. 2003)), this hypothesis does not agree with some researchers' observations of perithecia and ascospores for *M. nivale* (e.g. (Lees et al. 1995)), and also contrasts with the high genetic variability of this species that has been reported (Lees et al. 1995; Mahuku et al. 1998). Further implications and possible explanations for these observations are discussed below, in the context of the multi-species mating type analysis, as well as in Chapter 7.

To explore the surprising and contradictory observations made using the genomic information, isolates of *M. nivale* and *M. majus* were inoculated onto wheat straw on either PDA or water agar and incubated at 5, 10, 15, or 20 °C. Perithecia with ascospores were observed on several of the *M. majus* plates after only six weeks of incubation on water agar. Ascospores were produced both for single-spore isolates in self by self crosses and on plates that contained two
different isolates; however, there were no indications that the ascospores produced were the result of mating between the two isolates. In contrast, in most cases, there was a clear demarcation in the approximate centre of the wheat straw, suggesting that the perithecia were produced by one isolate alone (Figure 5.3). As all isolates used were derived from single conidia, and because the conidia of most Ascomycetes contain only a single nucleus (Webster and Weber 2007), these observations are in agreement with previous reports that *M. majus* is homothallic.

The putative ascospores identified in all of the apparent mating events were examined by measuring their size and recording their appearance. In all cases, these apparent ascospores matched the reported descriptions of ascospores from *M. nivale* sensu lato, and they were readily distinguished from conidia based on their "lumpier" appearance and their short length (10-17 µm (Booth 1971) compared to 15-33 µm (Glynn et al. 2005)) relative to conidia from the same isolate. When the conidial dimensions of eight *M. majus* isolates were recorded, the average size observed for the isolates tested was larger than that of the putative ascospores, confirming that the spores observed from the wheat-incubated colonies were truly ascospores. The genotypes of SA-derived cultures were compared to their "parent" isolate using SSR and ISSR. Seven additional isolates of *M. majus*, including one collected at the same time and from the same location as the parent isolate, were included in this analysis. The SSR and ISSR primers chosen for this experiment were those that produced different banding patterns for the *M. nivale* samples collected across three years (Chapter 4), and the results of this analysis were visualized by constructing a UPGMA tree with 100 bootstrap replicates.

Despite minor differences in the banding patterns produced by the eight SSR and ISSR primers, the ten SA isolates and their parent culture were more closely related to each other than to any of the other isolates included in this study. The four Ottawa isolates, 12043-12046, were
found in a single clade that also included one of the European isolates. In addition, the Atwood isolate 99061, which was collected on the same date and from the same location as isolate 99049, was not closely grouped with any of the 99049 isolates included. These results show that the ten SA cultures were indeed highly similar to their parent culture 99049. This result is in agreement with the hypothesis that these isolates are the product of homothallic mating by the sequenced *M. majus* isolate 99049, despite the apparent absence of the MAT1-1-1 gene in the genome of this or any other *Microdochium* isolate examined.

The temperature of incubation and the nutrient content of the media were found to play an important role in the production of perithecia, and possibly also the formation of ascospores. Ascospores were only observed among isolate crosses incubated on water agar, rather than on PDA; this suggests that ascospore production may be a direct response to nutrient starvation. In contrast, the production of ascospores was most frequent at 20 °C, a temperature at which these fungi grow rapidly in culture (Tronsmo et al. 2001). At 5 °C, a temperature that is below the optimum in-lab growth temperature but well within the range of temperatures at which these species cause disease (Smiley et al. 2005), ascospore production was not observed.

Together, these observations suggest that the production of ascospores in the field may be most common when these fungi are not actively causing disease symptoms. Although these fungi grow readily on both artificial media and plant tissue in the lab at 20 °C, in the field, these temperatures also favour the growth of several other pathogenic fungi that are active on the same hosts (Smiley et al. 2005). It is thus possible that ascospores and / or perithecia may provide an alternate mechanism for over-summering. In 1993, Koizuma (in Tronsmo et al. 2001) reported that perithecia may survive on plant matter on the soil for up to three months; however, to
investigate the role that ascospores may play as a reservoir of inoculum future infections, the longevity of ascospores under a variety of conditions should be investigated.

Based on the results observed from the in-lab mating experiments, which corroborate published reports that *M. majus* (and likely *M. nivale*) are homothallic, the MAT1 and flanking genes from several non-*Microdochium* species from the Xylariales were investigated. The goal of this experiment was to determine whether a putative MAT1-1-1 sequence could be identified in any of these species. If a Xylariales MAT1-1-1 could be identified, this sequence might facilitate the identification of MAT1-1-1 from *M. nivale* and / or *M. majus*. The synteny of the flanking genes was also investigated to determine whether the unusual layout observed in *M. majus* and *M. majus* was specific to this genus, or whether it is conserved among the Xylariales.

For these reasons, the MAT1 locus and flanking genes in seven non-*Microdochium* genomes were analysed. An additional three *Microdochium* genomes (*M. nivale* isolate 10106, *M. majus* isolate 10095, and *M. bolleyi* isolate 07020) that became available at this time were also included in this analysis. The "flanking region" in this study was expanded to include two additional genes, APC5 (anaphase-promoting complex) and COX13 (cytochrome c co-oxidase) to facilitate the identification of any further patterns of organization among the genes in this region.

In all thirteen of the Xylariales genomes studied, the four flanking genes and at least two HMG-box-containing sequences were identified. None of the genomes contained a putative match for MAT1-1-1. In the six *Microdochium* spp. genomes, the synteny of the MAT1 locus and the four flanking genes was identical; however, this relationship was only conserved in *Pestalotiopsis* sp. (Figure 5.2). In the other Xylariales genomes, the only trend observed was that APC5 and SLA2 were always in close proximity to each other, as were APN2 and COX13. In
*Daldinia eschscholtzii*, all four genes were in very close proximity with only 3.4 kb separating APN2 and SLA2; in *Annulohypoxylon stygium* and two of the three *Hypoxylon* sp. genomes (genomes "EC" and "CI"), this distance was greater than 10 kb, and in the third *Hypoxylon* sp. genome, the two groups of flanking genes were found on separate scaffolds. In these five cases, all of the putative HMG-containing sequences identified were also on separate scaffolds relative to any of the flanking genes. In *P. neglecta*, all four of the flanking genes and all of the HMG-box-containing sequences were found on separate contigs. In those species where the HMG-box containing sequences were distant from the flanking genes, it was not possible to determine which, if any, of the sequences identified might be MAT1-2-1. Turgeon and Yoder (2000) recognized the difficulty in distinguishing between MAT1-1-3 and MAT1-2-1, and suggested that the presence of the other MAT1-1 genes (MAT1-1-1 and MAT1-1-2) in "true" MAT1-1 strains may be a useful diagnostic. However, likely homologs for neither MAT1-1-1 nor MAT1-1-2 were found in any of the *Microdochium* spp. genomes studied.

The lack of synteny among the other Xylariales species relative to the *Microdochium* spp. and *Pestalotiopsis* sp. MAT1 loci and flanking regions led to a detailed investigation of all of the matches identified for all of the genomes studied. The search was performed to clarify whether any of the secondary matches for the flanking genes represented truncated or otherwise non-functional duplicate copies of the genes of interest. However, no putative duplicates were detected based on a comparison of these candidate sequences to the GenBank database (BLASTx).

Although few reports on sexual reproduction among the Xylariales are available, those species that do produce perithecia in the lab are reported to be homothallic (Rogers 1979). This general observation is in line with the proposed mode of sexual reproduction in *M. nivale* and *M.*
majus. The apparent lack of MAT1-1-1 in all of the Xylariales examined was thus unexpected. One possible explanation for the apparent discrepancy is that MAT1-1-1 truly is found in the genomes of all homothallic Xylariales, but that the sequence of this gene among this order is simply dissimilar from the MAT1-1-1 sequences previously reported, making it difficult to find with the tools and information that were used in these experiments. The proposed sequence divergence may be explained by the apparently basal relationship of the Xylariales to the other members of the Sordariomycetes (Schoch et al. 2009). The proposed dissimilarity may also explain the unusual and generally inconsistent arrangement of the flanking genes and putative MAT1 loci of these species relative to that of other Sordariomycetes. If this hypothesis is true, the yet-unidentified MAT1-1-1 gene is most likely found in a location that is physically distant from any of the genes that are "traditionally" identified as syntenic to the MAT1 cassette, because the regions 10 kb immediately up- and down-stream of the flanking genes were also searched, unsuccessfully, for sequences resembling MAT1-1-1. In some homothallic ascomycetes where the MAT genes are distant from one another, one gene may be surrounded by both of the flanking genes while the other is proximate to at least a partial sequence of one of the two genes (e.g. (Pöggeler et al. 2011)). However, neither complete nor truncated duplicate sequences for any of the four flanking genes were identified in the genomes of any of the Xylariales examined, which complicates future efforts to identify a MAT1-1-1 sequence among the Xylariales.

A more intriguing explanation is that the MAT1-1-1 sequence truly is absent in these genomes, and that mating in the Xylariales may not be under the exclusive control of the "canonical" MAT genes MAT1-1-1 and MAT1-2-1. This hypothesis is strengthened by recent reports of unconventional mating among other Ascomycota, including members of the genera
Glomerella (Chen et al. 2002; Menat et al. 2012; Rodriguez-Guerra et al. 2005; Vaillancourt et al. 2000) and Neurospora (Lin and Heitman 2007) In those Neurospora species displaying apparent homothallism but possessing only a single MAT gene, only MAT1-1-1 was identified (Lin and Heitman 2007). However, in G. graminicola, strains displaying a complex mixture of self- and inter-fertility and sterility that could not be readily classified as homo- or heterothallism apparently contained only MAT1-2-1 (Vaillancourt et al. 2000). This situation is similar to that observed for the Xylariales examined in this Chapter, especially M. nivale, which apparently possessed two distinct alleles of the putative MAT1-2-1 gene (exemplified by the genes from isolates 12262 and 10106 vs. that of 11037). At this stage, a complete explanation of the mating behaviour in M. nivale and M. majus is not possible.

The results presented in this Chapter underscore the importance of corroborating genetic data with field and laboratory observations of living samples, and vice versa. Apparently homothallic mating was observed among isolates of M. majus, which produced fertile ascospores, but the few isolates of M. nivale that produced perithecia failed to produce ascospores. An examination of genomic data for three isolates of M. nivale, two isolates of M. majus, and a total of eight other members of the Xylariales failed to reveal the presence of a second MAT gene, suggesting that either the nucleotide sequence of the MAT1-1-1 gene among the Xylariales is so different from that of other Sordariomycetes that it is unrecognizable, or that this gene may be missing altogether from this order, and that the control of mating in this group may be unlike that in other taxa. Further research is necessary to explore these possible explanations.
5.5 References for Chapter 5


<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer Name</th>
<th>Primer Sequence</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAT1-1-1</td>
<td>Falpha1</td>
<td>CGGTCAYGAGTATCTTCCTG</td>
<td>(Arie et al. 2000)</td>
</tr>
<tr>
<td></td>
<td>Falpha2</td>
<td>GATGTAGATGGAGGTCA</td>
<td>(Arie et al. 2000)</td>
</tr>
<tr>
<td>MAT1-2-1</td>
<td>ChHMG1</td>
<td>AAGGCNCCNCYGCCNATGAA</td>
<td>(Arie et al. 1997)</td>
</tr>
<tr>
<td></td>
<td>ChHMG2</td>
<td>CTNGGNATGATTTGTAAATTNGG</td>
<td>(Arie et al. 1997)</td>
</tr>
</tbody>
</table>
### Table 5.2 Isolates of *M. nivale* and *M. majus* tested with published mating-type primers

<table>
<thead>
<tr>
<th>Species</th>
<th>Isolate number</th>
<th>Geographic origin</th>
<th>Host plant origin</th>
<th>Primer(s) Tested*</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. majus</td>
<td>99061</td>
<td>NA</td>
<td>W</td>
<td>ChHMG, Fα</td>
</tr>
<tr>
<td></td>
<td>99049</td>
<td>NA</td>
<td>W</td>
<td>Fα</td>
</tr>
<tr>
<td></td>
<td>10082</td>
<td>UK</td>
<td>T</td>
<td>ChHMG, Fα</td>
</tr>
<tr>
<td></td>
<td>10083</td>
<td>UK</td>
<td>T</td>
<td>ChHMG, Fα</td>
</tr>
<tr>
<td></td>
<td>99046</td>
<td>NA</td>
<td>W</td>
<td>ChHMG, Fα</td>
</tr>
<tr>
<td></td>
<td>99006</td>
<td>NA</td>
<td>W</td>
<td>ChHMG, Fα</td>
</tr>
<tr>
<td>M. nivale</td>
<td>99069</td>
<td>NA</td>
<td>W</td>
<td>ChHMG, Fα</td>
</tr>
<tr>
<td></td>
<td>99007</td>
<td>NA</td>
<td>W</td>
<td>ChHMG, Fα</td>
</tr>
<tr>
<td></td>
<td>10086</td>
<td>NA</td>
<td>T</td>
<td>ChHMG, Fα</td>
</tr>
<tr>
<td></td>
<td>10084</td>
<td>UK</td>
<td>T</td>
<td>ChHMG</td>
</tr>
<tr>
<td></td>
<td>10086</td>
<td>NA</td>
<td>T</td>
<td>ChHMG</td>
</tr>
</tbody>
</table>

* The literature references for these primers are found in Table 5.1.
Table 5.3 List of species and GenBank accession numbers for genes used to design conserved MAT1-1-1 and MAT1-2-1 primers in *Microdochium* sp.

<table>
<thead>
<tr>
<th>Gene of interest</th>
<th>Genus and Species</th>
<th>GenBank Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Alternaria tenuissima</em></td>
<td></td>
<td>AY004675.1</td>
</tr>
<tr>
<td><em>Pleospora triglochinicola</em></td>
<td></td>
<td>AY335167.1</td>
</tr>
<tr>
<td><em>Pleospora eturmiuna</em></td>
<td></td>
<td>AY335176.1</td>
</tr>
<tr>
<td><em>Stemphylium majusculum</em></td>
<td></td>
<td>AY335174.1</td>
</tr>
<tr>
<td><em>Pleospora tarda</em></td>
<td></td>
<td>AY335164.1</td>
</tr>
<tr>
<td><em>Stemphylium callistephi</em></td>
<td></td>
<td>AY339863.1</td>
</tr>
<tr>
<td><em>Pleospora paludiscirpi</em></td>
<td></td>
<td>AY335177.1</td>
</tr>
<tr>
<td><em>Stemphylium solani</em></td>
<td></td>
<td>AY339855.1</td>
</tr>
<tr>
<td><em>Lewia infectoria</em></td>
<td></td>
<td>AB444188.1</td>
</tr>
<tr>
<td><em>Penicillium decumbens</em></td>
<td></td>
<td>HM067979.1</td>
</tr>
<tr>
<td><em>Mycosphaerella musicola</em></td>
<td></td>
<td>GU057991.1</td>
</tr>
<tr>
<td><em>Setosphaeria turcica</em></td>
<td></td>
<td>GU997138.1</td>
</tr>
<tr>
<td><em>Ophiostoma novo-ulmi</em> var. novo-ulmi</td>
<td></td>
<td>FJ858801.1</td>
</tr>
<tr>
<td><em>Diaporthe melonis</em></td>
<td></td>
<td>GQ250237.1</td>
</tr>
<tr>
<td><em>Paracoccidioides brasiliensis</em></td>
<td></td>
<td>GQ411379.1</td>
</tr>
<tr>
<td><em>Aspergillus parasiticus</em></td>
<td></td>
<td>EU357935.1</td>
</tr>
<tr>
<td><strong>MAT1-1-1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Rhynchosporium secalis</em></td>
<td></td>
<td>FJ382949.1</td>
</tr>
<tr>
<td><em>Mycosphaerella pini</em></td>
<td></td>
<td>DQ915450.1</td>
</tr>
<tr>
<td><em>Dothistroma pini</em></td>
<td></td>
<td>DQ915449.1</td>
</tr>
<tr>
<td><em>Cercospora zeina</em></td>
<td></td>
<td>DQ264748.1</td>
</tr>
<tr>
<td><em>Cercospora apii</em></td>
<td></td>
<td>DQ264736.1</td>
</tr>
<tr>
<td><em>Fusarium oxysporum</em></td>
<td></td>
<td>AY527423.1</td>
</tr>
<tr>
<td><em>Pleospora sp.</em></td>
<td></td>
<td>AY339862.1</td>
</tr>
<tr>
<td><em>Fusarium brasilicum</em></td>
<td></td>
<td>AY452907.1</td>
</tr>
<tr>
<td><em>Alternaria alternaria</em></td>
<td></td>
<td>AB468151.1</td>
</tr>
<tr>
<td><em>Ophiostoma novo-ulmi</em></td>
<td></td>
<td>EU163846.1</td>
</tr>
<tr>
<td><em>Verticillium dahliae</em></td>
<td></td>
<td>AB469828.1</td>
</tr>
<tr>
<td><em>Phaeosphaeria avenaria</em> f.sp. triticae</td>
<td></td>
<td>AY196993.1</td>
</tr>
<tr>
<td><em>Pyrenophora tritici-repentis</em></td>
<td></td>
<td>AM884611.1</td>
</tr>
<tr>
<td><em>Neotyphodium uncinatum</em></td>
<td></td>
<td>AB258373.1</td>
</tr>
<tr>
<td><em>Fusarium oxysporum</em> f.sp. lycopersei</td>
<td></td>
<td>AB011379.2</td>
</tr>
<tr>
<td><em>Fusarium oxysporum</em></td>
<td></td>
<td>AY527415.1</td>
</tr>
<tr>
<td><strong>MAT1-2-1</strong></td>
<td><em>Xanthodactylon flammeum</em></td>
<td>CAI59780</td>
</tr>
<tr>
<td>Gene of interest</td>
<td>Genus and Species</td>
<td>GenBank Accession Number</td>
</tr>
<tr>
<td>------------------------------------------------------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Ajellomyces capsulatus</td>
<td></td>
<td>ABO87595</td>
</tr>
<tr>
<td>Petromyces alliaceus</td>
<td></td>
<td>ACE74241</td>
</tr>
<tr>
<td>Aspergillus flavus</td>
<td></td>
<td>ACA51904</td>
</tr>
<tr>
<td>Neosartorya fischeri</td>
<td></td>
<td>XP_001263957</td>
</tr>
<tr>
<td>Paracoccidioides brasiliensis</td>
<td></td>
<td>EEH50039</td>
</tr>
<tr>
<td>Coccidioides immitis</td>
<td></td>
<td>XP_001246635</td>
</tr>
<tr>
<td>Paracoccidioides brasiliensis</td>
<td></td>
<td>ACV32366</td>
</tr>
<tr>
<td>Microsporum gypseum</td>
<td></td>
<td>ACS91132</td>
</tr>
<tr>
<td>Dothistroma pini</td>
<td></td>
<td>ABK91353</td>
</tr>
<tr>
<td>Magnaporthe grisea</td>
<td></td>
<td>BAC65090</td>
</tr>
<tr>
<td>Cercospora zeae-maydis</td>
<td></td>
<td>ABB83719</td>
</tr>
<tr>
<td>Mycosphaerella eumusae</td>
<td></td>
<td>ADB11112</td>
</tr>
<tr>
<td>Glomerella lindemuthiana</td>
<td></td>
<td>ABY84976</td>
</tr>
<tr>
<td>Ophiostoma novo-ulmi</td>
<td></td>
<td>AAX83067</td>
</tr>
<tr>
<td>Sordaria macrospora</td>
<td></td>
<td>CAA71624</td>
</tr>
<tr>
<td>Neurospora cerealis</td>
<td></td>
<td>AAL28011</td>
</tr>
</tbody>
</table>
Table 5.4 Primers designed to amplify mating-type (MAT1-1 and MAT1-2) and flanking genes in *Microdochium* spp.

<table>
<thead>
<tr>
<th>Target Gene</th>
<th>Sequence Reference</th>
<th>Primer Name</th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAT1-1</td>
<td>GenBank</td>
<td>MAT1_3149F</td>
<td>TCATRGCYTTYCGMWGTAAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MAT1_3644R</td>
<td>GTGGTCATGATGCNNTTC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MAT1_3132F</td>
<td>AACGSCTACATGCGCTT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MAT1_3303R</td>
<td>GATTGCAARGTCTAYTCTYTT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MAT1_86F</td>
<td>CGGCTTTATGCGTTTC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MAT1_891R</td>
<td>AAGCTATAACATGCGGCAAT</td>
</tr>
<tr>
<td>MAT1-2</td>
<td>GenBank</td>
<td>MAT2_3515F</td>
<td>AGTTCTGATGATGGACGCTCA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MAT2_4290R</td>
<td>CAAGGGAAARCRCMCTSGT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MAT2_1404F</td>
<td>CCTYCATYCTSTACGCAGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MAT2_1811R</td>
<td>TACASCGMCTTGYCOCGA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MAT2_488F</td>
<td>AATGCCTACATTTCTTGCGC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MAT2_650R</td>
<td>ACGGACTGTGCTTCCCTCCCG</td>
</tr>
<tr>
<td></td>
<td>Mm99049</td>
<td>Mn_MAT2_3347F</td>
<td>TCGCCGCTATCCCACT</td>
</tr>
<tr>
<td></td>
<td>Mn11037</td>
<td>Mn_MAT2_20F</td>
<td>AGCATGCGACGATGAGCCC</td>
</tr>
<tr>
<td></td>
<td>Mn11037,</td>
<td>Mn_MAT2_198F</td>
<td>CGAACGCGAGGRCGAAG</td>
</tr>
<tr>
<td></td>
<td>Mn12262,</td>
<td>Mic_MAT2_676R</td>
<td>AAGCTCTGRTCTTGCTG</td>
</tr>
<tr>
<td></td>
<td>Mn99049</td>
<td>Mn_lyase_838F</td>
<td>CGCCCAAGAACCTTCC</td>
</tr>
<tr>
<td></td>
<td>Mn11037</td>
<td>Mn_APN2_700R</td>
<td>GCCCAAGAACCTTCC</td>
</tr>
<tr>
<td>APN2</td>
<td></td>
<td>Mic_APN2_32R</td>
<td>GAACCTTCTCCCACCTATCAC</td>
</tr>
<tr>
<td></td>
<td>Mm99049</td>
<td>Mn_SLA2_23R</td>
<td>TTGTGCGGCCACGGTGCGG</td>
</tr>
<tr>
<td></td>
<td>Mm99049</td>
<td>Mn_SLA2_1156F</td>
<td>CTCAGGACACGGAGGGCAG</td>
</tr>
<tr>
<td>SLA2</td>
<td>Mn11037</td>
<td>Mn_SLA2_357F</td>
<td>CCGCAATGTCGGCAAACA</td>
</tr>
<tr>
<td></td>
<td>Mn11037,</td>
<td>Mic_SLA2_92F</td>
<td>GCATGCTGCGTGCGCATCCT</td>
</tr>
<tr>
<td></td>
<td>Mn12262,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mm99049</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

320
Table 5.5 Number of bands amplified by the ISSR and SSR PCR primers listed with a selection of *Microdochium majus* isolates including the single-ascospore-derived (AS) cultures derived from the parent isolate 99049.

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Primer tested</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>916061</th>
<th>924078</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BHY(AGC)&lt;sub&gt;5&lt;/sub&gt;</td>
<td>DD(CCA)&lt;sub&gt;5&lt;/sub&gt;</td>
<td>(CAC)&lt;sub&gt;5&lt;/sub&gt;</td>
<td>(CT)&lt;sub&gt;5&lt;/sub&gt;</td>
<td>(GC)&lt;sub&gt;5&lt;/sub&gt;</td>
<td>(GAT)&lt;sub&gt;6&lt;/sub&gt;</td>
<td>916061</td>
<td>924078</td>
<td></td>
</tr>
<tr>
<td>99049 - AS1*</td>
<td>8</td>
<td>7</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>99049 - AS2</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>99049 - AS3</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>99049 - AS4</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>99049 - AS5</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>99049 - AS6</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>99049 - AS7</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>99049 - AS8</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>99049 - AS9</td>
<td>8</td>
<td>7</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>99049 - AS10</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>99049 - parent</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>12043</td>
<td>8</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>12044</td>
<td>8</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>12045</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>12046</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>10098</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>10099</td>
<td>8</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>99061</td>
<td>4</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Total number</td>
<td>8</td>
<td>7</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>of possible</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bands</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* AS = single-ascospore strain of isolate 99049
Table 5.6 List of species and GenBank accession numbers for genes used to search for putative MAT1-1-1, MAT1-2-1, and flanking genes in *Microdochium* genomes by standalone tBLASTn

<table>
<thead>
<tr>
<th>Genus and Species</th>
<th>Gene Name</th>
<th>GenBank Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Alternaria tenuissima</em></td>
<td>MAT1-1-1</td>
<td>AY004675.1</td>
</tr>
<tr>
<td><em>Pleospora triglochincola</em></td>
<td>MAT1-1-1 and MAT1-2-1</td>
<td>AY335167.1</td>
</tr>
<tr>
<td><em>Pleospora eturmiuna</em></td>
<td>MAT1-1-1 and MAT1-2-1</td>
<td>AY335176.1</td>
</tr>
<tr>
<td><em>Stemphylium majusculum</em></td>
<td>MAT1-1-1 and MAT1-2-1</td>
<td>AY335174.1</td>
</tr>
<tr>
<td><em>Microsporum gypseum</em></td>
<td>SLA2</td>
<td>ACS91152.1</td>
</tr>
<tr>
<td><em>Microsporum gypseum</em></td>
<td>APN2</td>
<td>ACS91139.1</td>
</tr>
<tr>
<td><em>Grossmania clavigera</em></td>
<td>COX13</td>
<td>EFX05020.1</td>
</tr>
<tr>
<td><em>Sordaria macrospora</em></td>
<td>APC5</td>
<td>XM_003347060.1</td>
</tr>
</tbody>
</table>
**Table 5.7** List of predicted genes corresponding to the mating-type (MAT1-2-1) and flaking genes (cytoskeletal assembly protein SLA2 and DNA lyase APN2) in the *Microdochium* genomes studied. Comparisons were performed using standalone tBLASTn to query the gene of interest against the *Microdochium* spp. genomes listed.

<table>
<thead>
<tr>
<th>Gene of Interest</th>
<th>Genome</th>
<th>Putative Homolog</th>
<th>Sequence Length (bp)</th>
<th>Similarity to sequence used to identify putative homolog</th>
<th>Alignment length (nt)</th>
<th>Identity (%)</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLA2</td>
<td>Mm99049</td>
<td>g9065</td>
<td>3,144</td>
<td></td>
<td>3207</td>
<td>68.8</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>Mn11037</td>
<td>g10271</td>
<td>3,156</td>
<td></td>
<td>3219</td>
<td>68.1</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>Mn12262</td>
<td>g814</td>
<td>3,144</td>
<td></td>
<td>3207</td>
<td>68.8</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>Mn10095</td>
<td>g10807</td>
<td>3,144</td>
<td></td>
<td>3207</td>
<td>68.8</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>Mn10106</td>
<td>g6682</td>
<td>3,144</td>
<td></td>
<td>3207</td>
<td>68.8</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>Mb07020</td>
<td>g6892</td>
<td>3,450</td>
<td></td>
<td>3207</td>
<td>69.0</td>
<td>0.0</td>
</tr>
<tr>
<td>MAT1-2-1</td>
<td>Mm99049</td>
<td>g9066</td>
<td>1,860</td>
<td></td>
<td>210</td>
<td>34.3</td>
<td>2e-12</td>
</tr>
<tr>
<td></td>
<td>Mn11037</td>
<td>g10270</td>
<td>1,671</td>
<td></td>
<td>720</td>
<td>23.3</td>
<td>3e-11</td>
</tr>
<tr>
<td></td>
<td>Mn12262</td>
<td>g814</td>
<td>1,941</td>
<td></td>
<td>177</td>
<td>77</td>
<td>2e-06</td>
</tr>
<tr>
<td></td>
<td>Mn10095</td>
<td>g10806</td>
<td>1,722</td>
<td></td>
<td>141</td>
<td>34.0</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Mn10106</td>
<td>g6681</td>
<td>1,890</td>
<td></td>
<td>210</td>
<td>32.9</td>
<td>2e-08</td>
</tr>
<tr>
<td></td>
<td>Mb07020</td>
<td>g6891</td>
<td>1,815</td>
<td></td>
<td>222</td>
<td>37.8</td>
<td>3e-11</td>
</tr>
<tr>
<td>APN2</td>
<td>Mm99049</td>
<td>g9067</td>
<td>1,824</td>
<td></td>
<td>1791</td>
<td>46.9</td>
<td>2e-142</td>
</tr>
<tr>
<td></td>
<td>Mn11037</td>
<td>g10269</td>
<td>1,962</td>
<td></td>
<td>1917</td>
<td>46.3</td>
<td>7e-147</td>
</tr>
<tr>
<td></td>
<td>Mn12262</td>
<td>g816</td>
<td>1,866</td>
<td></td>
<td>1764</td>
<td>47.4</td>
<td>7e-142</td>
</tr>
<tr>
<td></td>
<td>Mn10095</td>
<td>g10805</td>
<td>1,839</td>
<td></td>
<td>1818</td>
<td>48.7</td>
<td>1e-150</td>
</tr>
<tr>
<td></td>
<td>Mn10106</td>
<td>g6680</td>
<td>2,253</td>
<td></td>
<td>1764</td>
<td>48.8</td>
<td>3e-150</td>
</tr>
<tr>
<td></td>
<td>Mb07020</td>
<td>g6890</td>
<td>1,833</td>
<td></td>
<td>1785</td>
<td>49.1</td>
<td>2e-156</td>
</tr>
</tbody>
</table>

* Based on coding sequence of predicted gene
† Based on tBLASTn search against a database consisting of the genome of interest
Table 5.8 Results of MAT1-2-1 amplification of *M. nivale* isolates with the primers Mn_MAT2_20F and Mn_MAT2_727R

<table>
<thead>
<tr>
<th>Host Plant</th>
<th>Collection location*</th>
<th>Number of isolates tested</th>
<th>Number of isolates amplified by MAT1-2-1 primers</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>P. pratensis</em></td>
<td>GTI (beside native green)</td>
<td>25</td>
<td>6</td>
</tr>
<tr>
<td><em>P. pratensis</em></td>
<td>GTI (beside road)</td>
<td>25</td>
<td>21</td>
</tr>
<tr>
<td><em>P. annua / A. stolonifera mix</em></td>
<td>GTI (native green)</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td><em>L. perenne</em></td>
<td>GTI (roadway)</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td><em>Triticum sp.</em></td>
<td>Ottawa Experimental Farm</td>
<td>23</td>
<td>13</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td><strong>92</strong></td>
<td><strong>54</strong></td>
</tr>
<tr>
<td>Experiment</td>
<td>Species</td>
<td>Host Plant</td>
<td>Isolate Number</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>------------</td>
<td>----------------</td>
</tr>
<tr>
<td>1</td>
<td><em>M. majus</em></td>
<td>wheat</td>
<td>99049</td>
</tr>
<tr>
<td></td>
<td><em>M. majus</em></td>
<td>wheat</td>
<td>10095</td>
</tr>
<tr>
<td></td>
<td><em>M. majus</em></td>
<td>wheat</td>
<td>11029</td>
</tr>
<tr>
<td></td>
<td><em>M. nivale</em></td>
<td>turfgrass</td>
<td>11037</td>
</tr>
<tr>
<td></td>
<td><em>M. nivale</em></td>
<td>turfgrass</td>
<td>12085</td>
</tr>
<tr>
<td></td>
<td><em>M. nivale</em></td>
<td>wheat</td>
<td>12262</td>
</tr>
<tr>
<td>2</td>
<td><em>M. majus</em></td>
<td>wheat</td>
<td>99049</td>
</tr>
<tr>
<td></td>
<td><em>M. majus</em></td>
<td>wheat</td>
<td>12045</td>
</tr>
<tr>
<td></td>
<td><em>M. majus</em></td>
<td>wheat</td>
<td>12166</td>
</tr>
</tbody>
</table>
Table 5.10 Summary of perithecial production in second experiment according to temperature of incubation and species included in each cross. Isolates were inoculated on sterilized wheat straw overlaid on water agar, and observations were performed after two months of incubation.

<table>
<thead>
<tr>
<th>Cross</th>
<th>Number of Plates</th>
<th>Percentage of plates with perithecia at each incubation temperature</th>
<th>Percentage of perithecia with ascospores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>5 °C</td>
<td>10 °C</td>
</tr>
<tr>
<td>M x M</td>
<td>10</td>
<td>0%</td>
<td>30%</td>
</tr>
<tr>
<td>M x N</td>
<td>20</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>N x N</td>
<td>15</td>
<td>0%</td>
<td>20%</td>
</tr>
</tbody>
</table>
Table 5.11 Species used for MAT-region synteny investigation. All species were members of the order Xylariales and family Xylariaceae

<table>
<thead>
<tr>
<th>Genus and Species</th>
<th>Genome Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Daldinia eschscholtzii</em></td>
<td>Joint Genome Institute</td>
</tr>
<tr>
<td><em>Annulohypoxylon stygium</em></td>
<td>courtesy of B. Xie, Fujian Agriculture and Forestry University, China</td>
</tr>
<tr>
<td><em>Hypoxylon</em> sp. (CO)</td>
<td>Joint Genome Institute</td>
</tr>
<tr>
<td><em>Hypoxylon</em> sp. (EC)</td>
<td>Joint Genome Institute</td>
</tr>
<tr>
<td><em>Pestalotiopsis neglecta</em></td>
<td>courtesy of K. Watanabe, Tamagawa University, Japan</td>
</tr>
<tr>
<td><em>Pestalotiopsis</em> sp.</td>
<td>courtesy of K. Watanabe, Tamagawa University, Japan</td>
</tr>
</tbody>
</table>
Table 5.12 BLASTx results for putative matches to flanking and mating genes observed in Xylariales genomes. BLAST searches were performed by querying the putative flanking genes (cytoskeletal assembly protein SLA2, DNA lyase APN2, anaphase-promoting complex protein APC5, and cytochrome oxidase COX) and mating type gene (MAT2) against the GenBank non-redundant database.

<table>
<thead>
<tr>
<th>Species</th>
<th>Query Identity</th>
<th>Query location in genome</th>
<th>Top Hit</th>
<th>Accession</th>
<th>%ID (\uparrow)</th>
<th>%Query (\downarrow)</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Daldinia eschscholtzii</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLA2</td>
<td>scaffold_13</td>
<td>putative cytoskeleton assembly control protein sla2 protein</td>
<td>EMR72087.1</td>
<td>88%</td>
<td>98%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold_19</td>
<td>arsenite resistance</td>
<td>EMR69908.1</td>
<td>97%</td>
<td>98%</td>
<td>2.00E-19</td>
<td></td>
</tr>
<tr>
<td>APN2</td>
<td>scaffold_13</td>
<td>Hypothetical protein (probable APN2)</td>
<td>XP_003660213.1</td>
<td>78%</td>
<td>92%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold_55</td>
<td>Hypothetical protein (probable cytosolic ARG-trna ligase)</td>
<td>XP_001907561.1</td>
<td>73%</td>
<td>99%</td>
<td>9.00E-33</td>
<td></td>
</tr>
<tr>
<td>MAT2</td>
<td>scaffold_38</td>
<td>HMG box transcriptional</td>
<td>ELA37846.1</td>
<td>80%</td>
<td>98%</td>
<td>2.00E-29</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold_58</td>
<td>HMG box protein</td>
<td>XP_963833.2</td>
<td>68%</td>
<td>96%</td>
<td>1.00E-26</td>
<td></td>
</tr>
<tr>
<td>APC5</td>
<td>scaffold_13</td>
<td>anaphase promoting protein</td>
<td>EMR72084.1</td>
<td>78%</td>
<td>99%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold_42</td>
<td>amidophosphoribosyltransferase</td>
<td>XP_002488025.1</td>
<td>64%</td>
<td>75%</td>
<td>8.00E-36</td>
<td></td>
</tr>
<tr>
<td>COX</td>
<td>scaffold_13</td>
<td>Hypothetical protein (probablecytochrome c-oxidase-like protein)</td>
<td>XP_003651352.1</td>
<td>60%</td>
<td>99%</td>
<td>3.00E-30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold_27</td>
<td>SMAC protein</td>
<td>XP_003348334.1</td>
<td>100%</td>
<td>98%</td>
<td>8.00E-33</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold_3</td>
<td>no matches found</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Annulohypoxylon stygium</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLA2</td>
<td>scaffold285</td>
<td>putative cytoskeleton assembly control protein sla2 protein</td>
<td>EMR72087.1</td>
<td>94%</td>
<td>91%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold96</td>
<td>Hypothetical protein (probable HET_domain containing protein)</td>
<td>EJT71176.1</td>
<td>31%</td>
<td>29%</td>
<td>3.00E-06</td>
<td></td>
</tr>
<tr>
<td>APN2</td>
<td>scaffold285</td>
<td>Hypothetical protein (probable APN2)</td>
<td>XP_003660213.1</td>
<td>78%</td>
<td>94%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold149</td>
<td>Hypothetical protein (probable wsc)</td>
<td>EJT81425.1</td>
<td>73%</td>
<td>49%</td>
<td>7.00E-12</td>
<td></td>
</tr>
</tbody>
</table>

328
<table>
<thead>
<tr>
<th>Species</th>
<th>Query Identity</th>
<th>Query location in genome</th>
<th>Top Hit</th>
<th>Accession</th>
<th>%ID†</th>
<th>%Query‡</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>domain-containing protein)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold71</td>
<td>nitrate reductse</td>
<td>ADY76216.1</td>
<td>84%</td>
<td>99%</td>
<td>8.00E-90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold76</td>
<td>nuclear export protein</td>
<td>EMR70615.1</td>
<td>78%</td>
<td>98%</td>
<td>2.00E-19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold28</td>
<td>Putative HMG box protein</td>
<td>EMR65965.1</td>
<td>82%</td>
<td>99%</td>
<td>3.00E-31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold60</td>
<td>Hypothetical protein (probable HMG-box containing protein)</td>
<td>XP_003856925.1</td>
<td>59%</td>
<td>99%</td>
<td>2.00E-28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold285</td>
<td>putative anaphase-promoting complex protein</td>
<td>EMR72084.1</td>
<td>78%</td>
<td>99%</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold285</td>
<td>putative cytochrome c oxidase subunit protein</td>
<td>EMR72081.1</td>
<td>66%</td>
<td>99%</td>
<td>1.00E-34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold_187</td>
<td>predicted protein (probable SLA2)</td>
<td>XP_003053294.1</td>
<td>100%</td>
<td>95%</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold_296</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold_17</td>
<td>Hypothetical protein (probable SLA2)</td>
<td>XP_003660213.1</td>
<td>83%</td>
<td>97%</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold_184</td>
<td>amine oxidase</td>
<td>EFQ31663.1</td>
<td>63%</td>
<td>89%</td>
<td>8.00E-13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold_107</td>
<td>Hypothetical protein (probable HMG-box containing protein)</td>
<td>ELR08424.1</td>
<td>64%</td>
<td>95%</td>
<td>1.00E-24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold_85</td>
<td>HMG box transcriptional</td>
<td>ELA37846.1</td>
<td>83%</td>
<td>98%</td>
<td>4.00E-31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold_187</td>
<td>putative anaphase-promoting complex protein</td>
<td>EMR72084.1</td>
<td>79%</td>
<td>97%</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold_17</td>
<td>Hypothetical protein (probable cytochrome c oxidase subunit protein)</td>
<td>XP_003651352.1</td>
<td>54%</td>
<td>99%</td>
<td>1.00E-31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold_80</td>
<td>predicted protein (probable SLA2)</td>
<td>XP_003053294.1</td>
<td>100%</td>
<td>95%</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold_118</td>
<td>Hypothetical protein (unknown function)</td>
<td>CCF45826.1</td>
<td>76%</td>
<td>98%</td>
<td>3.00E-10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold_80</td>
<td>Hypothetical protein (probable APN2)</td>
<td>XP_003660213.1</td>
<td>75%</td>
<td>98%</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scaffold_128</td>
<td>Hypothetical protein (probable HMG-box containing protein)</td>
<td>EGS23343.1</td>
<td>60%</td>
<td>96%</td>
<td>4.00E-30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td>Query Identity</td>
<td>Query location in genome</td>
<td>Top Hit</td>
<td>Accession</td>
<td>%ID†</td>
<td>%Query‡</td>
<td>E-value</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------</td>
<td>--------------------------</td>
<td>---------------------------------------------</td>
<td>----------------</td>
<td>------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>Hypoxylon sp. (CI)</td>
<td>scaffold_18</td>
<td>HMG box transcriptional</td>
<td>Top Hit</td>
<td>ELA37846.1</td>
<td>83%</td>
<td>98%</td>
<td>4.00E-31</td>
</tr>
<tr>
<td></td>
<td>scaffold_80</td>
<td>putative anaphase-promoting complex</td>
<td></td>
<td>EMR72084.1</td>
<td>79%</td>
<td>98%</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>scaffold_80</td>
<td>cytochrome c oxidase-like protein</td>
<td></td>
<td>EGS22713.1</td>
<td>88%</td>
<td>52%</td>
<td>1.00E-29</td>
</tr>
<tr>
<td>SLA2</td>
<td>scaffold_26</td>
<td>putative cytoskeleton assembly control protein sla2</td>
<td></td>
<td>EMR72087.1</td>
<td>84%</td>
<td>98%</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>scaffold_47</td>
<td>putative tropomyosin-I alpha chain protein</td>
<td></td>
<td>EMR63401.1</td>
<td>64%</td>
<td>99%</td>
<td>5.00E-23</td>
</tr>
<tr>
<td></td>
<td>scaffold_19</td>
<td>fatty acid synthetase alpha subunit</td>
<td></td>
<td>XP_001836417.1</td>
<td>48%</td>
<td>98%</td>
<td>1.00E-06</td>
</tr>
<tr>
<td>APN2</td>
<td>scaffold_26</td>
<td>hypothetical protein (probable APN2)</td>
<td></td>
<td>XP_001903014.1</td>
<td>69%</td>
<td>97%</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>scaffold_19</td>
<td>no matches found</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAT2</td>
<td>scaffold_25</td>
<td>putative HMG box protein</td>
<td></td>
<td>EMR65965.1</td>
<td>85%</td>
<td>99%</td>
<td>1.00E-31</td>
</tr>
<tr>
<td></td>
<td>scaffold_40</td>
<td>HMG box protein</td>
<td></td>
<td>EFX01060.1</td>
<td>56%</td>
<td>99%</td>
<td>6.00E-23</td>
</tr>
<tr>
<td></td>
<td>scaffold_13</td>
<td>predicted insulin-like growth factor 1</td>
<td></td>
<td>XP_004069709.1</td>
<td>42%</td>
<td>80%</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>scaffold_3</td>
<td>hypothetical protein</td>
<td></td>
<td>EMR61840.1</td>
<td>68%</td>
<td>84%</td>
<td>3.00E-15</td>
</tr>
<tr>
<td>APC5</td>
<td>scaffold_26</td>
<td>putative anaphase-promoting complex protein</td>
<td></td>
<td>EMR72084.1</td>
<td>76%</td>
<td>98%</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>scaffold_26</td>
<td>putative cytochrome c oxidase subunit protein</td>
<td></td>
<td>EMR72081.1</td>
<td>89%</td>
<td>59%</td>
<td>6.00E-30</td>
</tr>
<tr>
<td></td>
<td>scaffold_77</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold_10</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold_4</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold_48</td>
<td>putative protophyrinogen oxidase</td>
<td></td>
<td>EMR69994.1</td>
<td>68%</td>
<td>93%</td>
<td>0.007</td>
</tr>
<tr>
<td>Pestalotiopsis neglecta</td>
<td>SLA2</td>
<td>pn1057 putative cytoskeletal assembly control protein</td>
<td></td>
<td>EMR72087.1</td>
<td>87%</td>
<td>95%</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>pn3470</td>
<td>putative acetylcholinesterase precursor protein</td>
<td></td>
<td>EMR61857.1</td>
<td>8300%</td>
<td>98%</td>
<td>4.00E-10</td>
</tr>
<tr>
<td>APN2</td>
<td>pn2907</td>
<td>Hypothetical protein (probable APN2)</td>
<td></td>
<td>XP_001903014.1</td>
<td>58%</td>
<td>99%</td>
<td>0</td>
</tr>
<tr>
<td>Species</td>
<td>Query Identity</td>
<td>Location in genome</td>
<td>Top Hit</td>
<td>Accession</td>
<td>%ID†</td>
<td>%Query‡</td>
<td>E-value</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------</td>
<td>--------------------</td>
<td>-------------------------------------------------------</td>
<td>--------------------</td>
<td>------</td>
<td>---------</td>
<td>--------------</td>
</tr>
<tr>
<td>MAT2</td>
<td>pn2367</td>
<td>Hypothetical protein (unknown function)</td>
<td>ELA32918.1</td>
<td>56%</td>
<td>53%</td>
<td>4.00E-08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pn2059</td>
<td>No matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pn2175</td>
<td>Hypothetical protein (probable nitrate reductase)</td>
<td>XP_003653361.1</td>
<td>77%</td>
<td>81%</td>
<td>3.00E-59</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pn1637</td>
<td>Hypothetical protein</td>
<td>XP_001593912.1</td>
<td>61%</td>
<td>75%</td>
<td>5.00E-14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pn713</td>
<td>SIT4 phosphate-associated protein</td>
<td>EGX95517.1</td>
<td>100%</td>
<td>51%</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pn1653</td>
<td>Hypothetical protein (probable pyruvate decarboxylase)</td>
<td>XP_002838525.1</td>
<td>44%</td>
<td>46%</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pn2098</td>
<td>No matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>APC5</td>
<td>pn1653</td>
<td>Anaphase promoting complex protein</td>
<td>EFQ29496.1</td>
<td>69%</td>
<td>95%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pn2907</td>
<td>Putative cytochrome c oxidase</td>
<td>EMR72081.1</td>
<td>91%</td>
<td>64%</td>
<td>3.00E-21</td>
<td></td>
</tr>
<tr>
<td>SLA2</td>
<td>scaffold110</td>
<td>Transmembrane actin binding-like protein</td>
<td>EGR50054.1</td>
<td>93%</td>
<td>96%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold44</td>
<td>Putative arsenite resistance protein</td>
<td>EMR69908.1</td>
<td>98%</td>
<td>98%</td>
<td>3.00E-10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold110</td>
<td>Hypothetical protein (putative APN2)</td>
<td>XP_001903014.1</td>
<td>66%</td>
<td>96%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold52</td>
<td>Hypothetical protein (unknown function)</td>
<td>XP_001903388.1</td>
<td>54%</td>
<td>99%</td>
<td>9.00E-59</td>
<td></td>
</tr>
<tr>
<td>Pestalotiopsis sp.</td>
<td>scaffold33</td>
<td>Transmembrane protein</td>
<td>ELA33651.1</td>
<td>67%</td>
<td>83%</td>
<td>6.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold716</td>
<td>Putative ochratoxin a non-ribosomal peptide synthetase protein</td>
<td>EMR70362.1</td>
<td>72%</td>
<td>98%</td>
<td>5.00E-20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold628</td>
<td>Hypothetical protein (probable mitochondrial carrier protein)</td>
<td>EFQ32338.1</td>
<td>95%</td>
<td>98%</td>
<td>4.00E-16</td>
<td></td>
</tr>
<tr>
<td>MAT2</td>
<td>scaffold318</td>
<td>HMG box transcriptional</td>
<td>ELA37846.1</td>
<td>86%</td>
<td>98%</td>
<td>4.00E-32</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold947</td>
<td>Hypothetical protein</td>
<td>XP_001905098.1</td>
<td>58%</td>
<td>95%</td>
<td>2.00E-21</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold110</td>
<td>Hypothetical protein (HMG-box containing protein)</td>
<td>XP_003051581.1</td>
<td>37%</td>
<td>97%</td>
<td>2.00E-10</td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td>Query Identity</td>
<td>Query location in genome</td>
<td>Top Hit</td>
<td>Accession</td>
<td>%ID†</td>
<td>%Query‡</td>
<td>E-value</td>
</tr>
<tr>
<td>---------</td>
<td>----------------</td>
<td>--------------------------</td>
<td>---------</td>
<td>--------------</td>
<td>------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>APC5</td>
<td>scaffold110</td>
<td>putative anaphase promoting complex protein</td>
<td>EMR72084.1</td>
<td>76%</td>
<td>96%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold99</td>
<td>HET domain containing protein</td>
<td>EIW82553.1</td>
<td>56%</td>
<td>84%</td>
<td>5.00E-26</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold443</td>
<td>No matches found</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COX</td>
<td>scaffold110</td>
<td>Putative cytochrome c oxidase subunit</td>
<td>EMR72081.1</td>
<td>90%</td>
<td>99%</td>
<td>2.00E-38</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold259</td>
<td>Putative c6 transcription factor protein</td>
<td>EMR66535.1</td>
<td>82%</td>
<td>98%</td>
<td>2.00E-15</td>
<td></td>
</tr>
<tr>
<td>SLA2</td>
<td>17551</td>
<td>putative cytoskeleton assembly control protein sla2</td>
<td>EMR72087.1</td>
<td>86%</td>
<td>97%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>17551</td>
<td>hypothetical protein (probable APN2)</td>
<td>XP_003660213.1</td>
<td>82%</td>
<td>94%</td>
<td>5.00E-177</td>
<td></td>
</tr>
<tr>
<td></td>
<td>17031</td>
<td>hypothetical protein (unknown function)</td>
<td>XP_003046908.1</td>
<td>69%</td>
<td>92%</td>
<td>4.00E-06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16665</td>
<td>hypothetical protein (unknown function)</td>
<td>YP_001657196.1</td>
<td>42%</td>
<td>60%</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td>M. majus 99049</td>
<td>17541</td>
<td>HMG box transcriptional</td>
<td>ELA37846.1</td>
<td>81%</td>
<td>98%</td>
<td>8.00E-30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>17527</td>
<td>hypothetical protein (HMG box protein)</td>
<td>EGR50539.1</td>
<td>40%</td>
<td>94%</td>
<td>5.00E-13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>17484</td>
<td>predicted protein (putative Ku-domain containing protein)</td>
<td>XP_004345286.1</td>
<td>38%</td>
<td>53%</td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>17551</td>
<td>HMG box protein</td>
<td>EFQ28928.1</td>
<td>90%</td>
<td>98%</td>
<td>8.00E-34</td>
<td></td>
</tr>
<tr>
<td>MAT2</td>
<td>17551</td>
<td>putative anaphase-promoting complex protein</td>
<td>EMR72084.1</td>
<td>73%</td>
<td>99%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>APC5</td>
<td>17577</td>
<td>no matches found</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>232</td>
<td>hypothetical protein (putative kinesin light chain)</td>
<td>EMD64013.1</td>
<td>78%</td>
<td>97%</td>
<td>0.042</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16833</td>
<td>hypothetical protein (putative superkiller protein)</td>
<td>EGY17196.1</td>
<td>64%</td>
<td>98%</td>
<td>2.00E-14</td>
<td></td>
</tr>
<tr>
<td>COX</td>
<td>17551</td>
<td>predicted protein (probable cytochrome</td>
<td>EEH23565.1</td>
<td>53%</td>
<td>63%</td>
<td>2.00E-25</td>
<td></td>
</tr>
</tbody>
</table>

332
<table>
<thead>
<tr>
<th>Species</th>
<th>Query Identity</th>
<th>Query location in genome</th>
<th>Top Hit</th>
<th>Accession</th>
<th>%ID†</th>
<th>%Query‡</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>c protein)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17568</td>
<td>no matches</td>
<td></td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16727</td>
<td>immunoglobin lambda light chain</td>
<td>ABA70877.1</td>
<td>34%</td>
<td>89%</td>
<td>1.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17604</td>
<td>cytochrome c oxidase subunit</td>
<td>EKG11568.1</td>
<td>43%</td>
<td>81%</td>
<td>1.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17575</td>
<td>no matches</td>
<td></td>
<td>no matches found</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLA2</td>
<td>scaffold11</td>
<td>putative cytoskeletal assembly control protein sla2 protein</td>
<td>EMR72087.1</td>
<td>88%</td>
<td>97%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold11</td>
<td>hypothetical protein (probable DNA lyase)</td>
<td>XP_003660213.1</td>
<td>79%</td>
<td>97%</td>
<td>1.00E-176</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold469</td>
<td>hypothetical protein (unknown function)</td>
<td>XP_003046908.1</td>
<td>69%</td>
<td>92%</td>
<td>4.00E-06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold45</td>
<td>hypothetical protein (unknown function)</td>
<td>YP_001657196.1</td>
<td>42%</td>
<td>60%</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>M. majus 10095</td>
<td>scaffold188</td>
<td>HMG box transcriptional</td>
<td>ELA37846.1</td>
<td>80%</td>
<td>81%</td>
<td>3.00E-23</td>
<td></td>
</tr>
<tr>
<td>APN2</td>
<td>scaffold11</td>
<td>putative HMG box (probable MAT1-2-1)</td>
<td>EMR65965.1</td>
<td>44%</td>
<td>75%</td>
<td>2.00E-06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold25</td>
<td>predicted protein (probable HMG box protein)</td>
<td>EGR50539.1</td>
<td>40%</td>
<td>94%</td>
<td>5.00E-13</td>
<td></td>
</tr>
<tr>
<td>MAT2</td>
<td>scaffold11</td>
<td>putative anaphase-promoting complex protein</td>
<td>EMR72084.1</td>
<td>73%</td>
<td>99%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold112</td>
<td>hypothetical protein (probable superkiller protein)</td>
<td>EGY17196.1</td>
<td>64%</td>
<td>98%</td>
<td>2.00E-14</td>
<td></td>
</tr>
<tr>
<td>APC5</td>
<td>scaffold11</td>
<td>predicted protein (probable cytochrome c oxidase)</td>
<td>EEH23565.1</td>
<td>53%</td>
<td>63%</td>
<td>2.00E-25</td>
<td></td>
</tr>
<tr>
<td>COX</td>
<td>scaffold259</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold40</td>
<td>immunoglobulin lambda light chain variable region</td>
<td>ABA70877.1</td>
<td>33%</td>
<td>93%</td>
<td>0.94</td>
<td></td>
</tr>
</tbody>
</table>

333
<table>
<thead>
<tr>
<th>Species</th>
<th>Query Identity</th>
<th>Query location in genome</th>
<th>Top Hit</th>
<th>Accession</th>
<th>%ID†</th>
<th>%Query‡</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. nivale</td>
<td>SLA2</td>
<td>scaffold302</td>
<td>cytochrome c oxidase</td>
<td>EKG11568.1</td>
<td>43%</td>
<td>81%</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>scaffold57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cytochrome c oxidase</td>
<td>EKG11568.1</td>
<td>43%</td>
<td>81%</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. nivale 12262</td>
<td>APN2</td>
<td>21826</td>
<td>putative cytoskeleton assembly control protein sla2</td>
<td>EMR72087.1</td>
<td>86%</td>
<td>98%</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22728</td>
<td>hypothetical protein (unknown function)</td>
<td>XP_003715561.1</td>
<td>69%</td>
<td>98%</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22786</td>
<td>uncharacterized protein</td>
<td>XP_003562413.1</td>
<td>65%</td>
<td>63%</td>
<td>4.2</td>
</tr>
<tr>
<td>M. nivale 12262</td>
<td>MAT2</td>
<td>21826</td>
<td>hypothetical protein (probable DNA lyase)</td>
<td>XP_003660213.1</td>
<td>65%</td>
<td>99%</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22484</td>
<td>ap endonuclease</td>
<td>XP_002842988.1</td>
<td>46%</td>
<td>93%</td>
<td>8.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22821</td>
<td>extracellular serine-rich protein</td>
<td>ELA33596.1</td>
<td>51%</td>
<td>45%</td>
<td>7.00E-05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21989</td>
<td>HMG box transcriptional</td>
<td>ELA37846.1</td>
<td>81%</td>
<td>98%</td>
<td>8.00E-30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22612</td>
<td>predicted protein (probable HMG box protein)</td>
<td>EGR50539.1</td>
<td>43%</td>
<td>94%</td>
<td>2.00E-14</td>
</tr>
<tr>
<td></td>
<td>APC5</td>
<td>21826</td>
<td>putative anaphase-promoting complex protein</td>
<td>EMR72084.1</td>
<td>74%</td>
<td>99%</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22417</td>
<td>hypothetical protein (possibly related to APC5)</td>
<td>ENH79118.1</td>
<td>65%</td>
<td>89%</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>COX</td>
<td>21826</td>
<td>hypothetical protein (probable cytochrome c oxidase)</td>
<td>XP_003347107.1</td>
<td>79%</td>
<td>55%</td>
<td>2.00E-26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22484</td>
<td>hypothetical protein (probable monoxygenase)</td>
<td>EME78310.1</td>
<td>65%</td>
<td>41%</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22715</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>22664</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>22792</td>
<td>hypothetical protein (probable phosphatidylinositol 4-kinase)</td>
<td>XP_001551465.1</td>
<td>100%</td>
<td>98%</td>
<td>2.00E-17</td>
</tr>
<tr>
<td>M. nivale 11037</td>
<td>SLA2</td>
<td>11541</td>
<td>putative cytoskeletal assembly control</td>
<td>EMR72087.1</td>
<td>88%</td>
<td>98%</td>
<td>0</td>
</tr>
<tr>
<td>Species</td>
<td>Query Identity</td>
<td>Query location in genome</td>
<td>Top Hit</td>
<td>Accession</td>
<td>%ID†</td>
<td>%Query‡</td>
<td>E-value</td>
</tr>
<tr>
<td>---------</td>
<td>----------------</td>
<td>--------------------------</td>
<td>---------</td>
<td>-----------</td>
<td>-------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>APN2</td>
<td>putative DNA lyase apn2 protein</td>
<td>CCD57793.1</td>
<td>53%</td>
<td>98%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10889</td>
<td>hypothetical protein</td>
<td>EIW83489.1</td>
<td>34%</td>
<td>8%</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11277</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAT2</td>
<td>HMG box transcriptional</td>
<td>ELA37846.1</td>
<td>80%</td>
<td>98%</td>
<td>3.00E-27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11541</td>
<td>putative mating type protein MAT-2</td>
<td>CAD59615.3</td>
<td>32%</td>
<td>99%</td>
<td>6.00E-08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>APC5</td>
<td>putative anaphase-promoting complex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>11541</td>
<td>hypothetical protein (probable DNA lyase)</td>
<td>XP_003347107.1</td>
<td>55%</td>
<td>99%</td>
<td>1.00E-28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>COX</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10911</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>11037</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>11448</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>11597</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M. nivale 10106</td>
<td>scaffold207</td>
<td>EMR72087.1</td>
<td>86%</td>
<td>98%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SLA2</td>
<td>putative cytoskeletal assembly control protein sla2 protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold48</td>
<td>hypothetical protein (unknown function)</td>
<td>XP_003715561.1</td>
<td>69%</td>
<td>98%</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold9</td>
<td>predicted protein (unknown function)</td>
<td>XP_003562413.1</td>
<td>65%</td>
<td>63%</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold207</td>
<td>hypothetical protein (probable DNA lyase)</td>
<td>XP_003660213.1</td>
<td>65%</td>
<td>99%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold639</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold320</td>
<td>extracellular serine-rich protein</td>
<td>ELA33596.1</td>
<td>51%</td>
<td>45%</td>
<td>7.00E-05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAT2</td>
<td>HMG box transcriptional</td>
<td>ELA37846.1</td>
<td>79%</td>
<td>76%</td>
<td>6.00E-21</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold207</td>
<td>putative HMG protein (probable MAT1-2-1)</td>
<td>EMR65965.1</td>
<td>41%</td>
<td>77%</td>
<td>3.00E-06</td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td>Query Identity</td>
<td>Query location in genome</td>
<td>Top Hit</td>
<td>Accession</td>
<td>%ID†</td>
<td>%Query‡</td>
<td>E-value</td>
</tr>
<tr>
<td>---------</td>
<td>----------------</td>
<td>--------------------------</td>
<td>---------</td>
<td>---------------</td>
<td>------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>APC5</td>
<td>scaffold203</td>
<td>predicted protein (probable HMG box protein)</td>
<td>EGR50539.1</td>
<td>43%</td>
<td>94%</td>
<td>2.00E-14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold207</td>
<td>putative anaphase-promoting complex protein</td>
<td>EMR72084.1</td>
<td>74%</td>
<td>99%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold165</td>
<td>hypothetical protein (unknown function)</td>
<td>ENH79118.1</td>
<td>65%</td>
<td>89%</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold207</td>
<td>hypothetical protein (probable cytochrome c oxidase)</td>
<td>XP_964326.1</td>
<td>73%</td>
<td>27%</td>
<td>2.00E-25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold543</td>
<td>dimethylamino monoxygenase</td>
<td>XP_001267756.1</td>
<td>64%</td>
<td>89%</td>
<td>5.00E-42</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold641</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold91</td>
<td>hypothetical protein (unknown function)</td>
<td>XP_001551465.1</td>
<td>97%</td>
<td>98%</td>
<td>2.00E-17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold327</td>
<td>hypothetical protein (probable phosphatidylinositol 4-kinase)</td>
<td>XP_001551465.1</td>
<td>97%</td>
<td>98%</td>
<td>2.00E-17</td>
<td></td>
</tr>
<tr>
<td>SLA2</td>
<td>scaffold41</td>
<td>putative cytoskeletal assembly control protein sla2 protein</td>
<td>EMR72087.1</td>
<td>85%</td>
<td>97%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold41</td>
<td>hypothetical protein (probable DNA lyase)</td>
<td>XP_003651351.1</td>
<td>84%</td>
<td>96%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold14</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold28</td>
<td>predicted low-quality protein</td>
<td>XP_003509871.1</td>
<td>54%</td>
<td>59%</td>
<td>2.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold7</td>
<td>putative hmg box protein</td>
<td>EMR65965.1</td>
<td>88%</td>
<td>98%</td>
<td>7.00E-32</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold41</td>
<td>MAT1-1-3</td>
<td>EKJ71585.1</td>
<td>38%</td>
<td>80%</td>
<td>8.00E-09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scaffold14</td>
<td>HMG box protein</td>
<td>EFQ28928.1</td>
<td>93%</td>
<td>82%</td>
<td>2.00E-41</td>
<td></td>
</tr>
<tr>
<td>MAT2</td>
<td>scaffold41</td>
<td>putative anaphase-promoting complex protein</td>
<td>EMR72084.1</td>
<td>74%</td>
<td>99%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>APC5</td>
<td>scaffold41</td>
<td>hypothetical protein (probable cytochrome c oxidase)</td>
<td>XP_002792088.1</td>
<td>82%</td>
<td>25%</td>
<td>5.00E-25</td>
<td></td>
</tr>
<tr>
<td>COX</td>
<td>scaffold41</td>
<td>hypothetical protein (probable cytochrome c oxidase)</td>
<td>XP_002792088.1</td>
<td>82%</td>
<td>25%</td>
<td>5.00E-25</td>
<td></td>
</tr>
</tbody>
</table>

*M. bolleyi* 07020
<table>
<thead>
<tr>
<th>Species</th>
<th>Query Identity</th>
<th>Query location in genome</th>
<th>Top Hit</th>
<th>Accession</th>
<th>%ID†</th>
<th>%Query‡</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>scaffold43</td>
<td>nonribosomal peptide synthase, putative</td>
<td></td>
<td>XP_003016838.1</td>
<td>61%</td>
<td>85%</td>
<td></td>
<td>2.00E-06</td>
</tr>
<tr>
<td>scaffold14</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>scaffold7</td>
<td>no matches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* When the top hit was a hypothetical or putative protein, this hit was in turn searched against the BLAST database to determine a tentative identity whenever possible.
† Percent identity (i.e. percent of nucleotides that were identical between the two sequences)
‡ Percent of query sequence that was represented in the match
**Figure 5.1** Orientation of the MAT1 region and flanking genes in the Sordariomycete species *Neurospora crassa* (Butler et al. 2004), *Giberella zeae* (Yun et al. 2000), *Botrytis cinerea*, and *Sclerotinia sclerotiorum* (Amselem et al. 2001). Diagram is not to scale. A vertical bar extending over the MAT1 locus indicates that this species is heterothallic and that the gene(s) located on the parallel bars are interchangeable in the two mating types.
Figure 5.2 Orientation and synteny of the putative MAT1 region and the flanking genes APC5, SLA2, APN2, and COX13 in several species of Xylariales, including *Microdochium* sp. Diagrams are not to scale. A double slash (//) indicates a long (>10 kb) distance between putative genes found on the same scaffold, and a vertical bar (|) indicates that the gene(s) that follow was / were found on a different scaffold.
Figure 5.3 Single segment of wheat straw inoculated with *M. majus* isolate 99049 and *M. nivale* isolate 99077 incubated on water agar at 20 °C for approximately two months. Note the production of perithecia on the side closest to *M. majus* (A) but not *M. nivale* (B). 10x magnification.
Figure 5.4 Perithecium of *M. majus* isolate 99049 at (A) 40 x and (B) 100x magnification.

Perithecia depicted were observed after two months of incubation on wheat straw on water agar at 20 °C
Figure 5.5 Ascospore (centre) produced by *M. majus* isolate 99049 at 400x magnification. Spore depicted were observed after two months of incubation on wheat straw on water agar at 20 °C.
**Figure 5.6** Bootstrapped UPGMA tree depicting the relationships between ten single-ascospore cultures derived from *M. majus* isolate 99049 relative to their parent culture and DNA from seven other *M. majus* isolates, including one isolate collected from the same location on the same date as the parent culture (99061), two isolates from Europe (10098 and 10099), and four cultures collected from the same wheat field on the same date (12043-12046). The horizontal bar represents 10% sequence divergence. Bootstrap values are out of 100.
Appendices for Chapter 5

Appendix 5.1 Alignment of MAT1-1-1 sequences collected from GenBank for primer design

<table>
<thead>
<tr>
<th>RefSeq</th>
<th>GenBank Accession</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AM983455.1Achr</td>
<td>FJ858801.10n-u</td>
<td>GCTGACTGACAGCGGCTGCATTGAGTTTTAAAAAGCAATGAGGAGTCTCGGTAGAGGAATTGGAAGAATAATGCGAGGTTGATCTTCGGTGGTGGTGGTGACCACATCTTTTTCCTCTATGCAAAAAAGGATGGGATACTTTGTTAAATTTTGTATTTCAATCTATGCCATATACATTTCCATCTCTAAAATATGT</td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td></td>
<td>-----------------------------------CATAATACACCAGTTTGCATAATAATGACAGGTCCAACAACAAATCATGCTCGAGGCTATTAATAGCGAGGTTGATCTTCGGTGGTGGTGGTGACCACATCTTTTTCCTCTATGCAAAAAAGGATGGGATACTTTGTTAAATTTTGTATTTCAATCTATGCCATATACATTTCCATCTCTAAAATATGT</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>FJ858801.10n-u</td>
<td>TTAATACGCCAAGCCCTTTTGGTGACTCTTCGACATTCGATCTCCATCGCCATCAGGGATAATCAGGAGGTTGATCTTCGGTGGTGGTGGTGACCACATCTTTTTCCTCTATGCAAAAAAGGATGGGATACTTTGTTAAATTTTGTATTTCAATCTATGCCATATACATTTCCATCTCTAAAATATGT</td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td></td>
<td>-----------------------------------CATAATACACCAGTTTGCATAATAATGACAGGTCCAACAACAAATCATGCTCGAGGCTATTAATAGCGAGGTTGATCTTCGGTGGTGGTGGTGACCACATCTTTTTCCTCTATGCAAAAAAGGATGGGATACTTTGTTAAATTTTGTATTTCAATCTATGCCATATACATTTCCATCTCTAAAATATGT</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>FJ858801.10n-u</td>
<td>TAAACACAGAATAGCCACCAATATGGCCACCAACAATGTCTGTTTTCTCTTTCTCTGGCA</td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td></td>
<td>-----------------------------------AGTTAGTATATGTTCACGCTCCTCAGGA</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>FJ858801.10n-u</td>
<td>TGAGAGCTGTCAATTTGACTATTCTTTAGTCCAGACATGACAGGTCCAACAACAAATCATGCTCGAGGCTATTAATAGCGAGGTTGATCTTCGGTGGTGGTGGTGACCACATCTTTTTCCTCTATGCAAAAAAGGATGGGATACTTTGTTAAATTTTGTATTTCAATCTATGCCATATACATTTCCATCTCTAAAATATGT</td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td></td>
<td>-----------------------------------TGTACCTCATAAAAGTTC</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>FJ858801.10n-u</td>
<td>TATCTTTGTCATGCTGACTAATGGGGCTGGCCATTTATTATCTGTTATTTGTACGCA</td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td></td>
<td>-----------------------------------TCTCCTGTAAGCAGTCAGTGCTCAGGAAACTCCTTTTCTCTCTCAGTTCCTAGGG</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>FJ858801.10n-u</td>
<td>TCAACAGAATAGCCACCAATATGGCCACCAACAATGTCTGTTTTCTCTTTCTCTGGCA</td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td></td>
<td>-----------------------------------AGTTAGTATATGTTCACGCTCCTCAGGA</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>FJ858801.10n-u</td>
<td>AGGCTCTGTGGACGAACTGGAAAAAGTAGACAGCGAAGTGGAAAAAGTGCAGCATTAAGA</td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td></td>
<td>-----------------------------------G---TCTATAAAATGTTCACGAGGGGTACCTCCTCTGACATTCAATCAGATGAGGAGTCTTGATCTCGGTGAGGTTGATCTTCGGTGGTGGTGGTGACCACATCTTTTTCCTCTATGCAAAAAAGGATGGGATACTTTGTTAAATTTTGTATTTCAATCTATGCCATATACATTTCCATCTCTAAAATATGT</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>FJ858801.10n-u</td>
<td>CGTCGGACAACAAATGCAACGTAAGCAACAAAAACATGACAGCGAAGTGGAAAAAGTGCAGCATTAAGA</td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td></td>
<td>-----------------------------------G---TCTATAAAATGTTCACGAGGGGTACCTCCTCTGACATTCAATCAGATGAGGAGTCTTGATCTCGGTGAGGTTGATCTTCGGTGGTGGTGGTGACCACATCTTTTTCCTCTATGCAAAAAAGGATGGGATACTTTGTTAAATTTTGTATTTCAATCTATGCCATATACATTTCCATCTCTAAAATATGT</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>FJ858801.10n-u</td>
<td>CGTCGGACAACAAATGCAACGTAAGCAACAAAAACATGACAGCGAAGTGGAAAAAGTGCAGCATTAAGA</td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td></td>
<td>-----------------------------------G---TCTATAAAATGTTCACGAGGGGTACCTCCTCTGACATTCAATCAGATGAGGAGTCTTGATCTCGGTGAGGTTGATCTTCGGTGGTGGTGGTGACCACATCTTTTTCCTCTATGCAAAAAAGGATGGGATACTTTGTTAAATTTTGTATTTCAATCTATGCCATATACATTTCCATCTCTAAAATATGT</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>FJ858801.10n-u</td>
<td>CGTCGGACAACAAATGCAACGTAAGCAACAAAAACATGACAGCGAAGTGGAAAAAGTGCAGCATTAAGA</td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td></td>
<td>-----------------------------------G---TCTATAAAATGTTCACGAGGGGTACCTCCTCTGACATTCAATCAGATGAGGAGTCTTGATCTCGGTGAGGTTGATCTTCGGTGGTGGTGGTGACCACATCTTTTTCCTCTATGCAAAAAAGGATGGGATACTTTGTTAAATTTTGTATTTCAATCTATGCCATATACATTTCCATCTCTAAAATATGT</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>FJ858801.10n-u</td>
<td>CGTCGGACAACAAATGCAACGTAAGCAACAAAAACATGACAGCGAAGTGGAAAAAGTGCAGCATTAAGA</td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td></td>
<td>-----------------------------------G---TCTATAAAATGTTCACGAGGGGTACCTCCTCTGACATTCAATCAGATGAGGAGTCTTGATCTCGGTGAGGTTGATCTTCGGTGGTGGTGGTGACCACATCTTTTTCCTCTATGCAAAAAAGGATGGGATACTTTGTTAAATTTTGTATTTCAATCTATGCCATATACATTTCCATCTCTAAAATATGT</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>FJ858801.10n-u</td>
<td>CGTCGGACAACAAATGCAACGTAAGCAACAAAAACATGACAGCGAAGTGGAAAAAGTGCAGCATTAAGA</td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td></td>
<td>-----------------------------------G---TCTATAAAATGTTCACGAGGGGTACCTCCTCTGACATTCAATCAGATGAGGAGTCTTGATCTCGGTGAGGTTGATCTTCGGTGGTGGTGGTGACCACATCTTTTTCCTCTATGCAAAAAAGGATGGGATACTTTGTTAAATTTTGTATTTCAATCTATGCCATATACATTTCCATCTCTAAAATATGT</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>FJ858801.10n-u</td>
<td>CGTCGGACAACAAATGCAACGTAAGCAACAAAAACATGACAGCGAAGTGGAAAAAGTGCAGCATTAAGA</td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td></td>
<td>-----------------------------------G---TCTATAAAATGTTCACGAGGGGTACCTCCTCTGACATTCAATCAGATGAGGAGTCTTGATCTCGGTGAGGTTGATCTTCGGTGGTGGTGGTGACCACATCTTTTTCCTCTATGCAAAAAAGGATGGGATACTTTGTTAAATTTTGTATTTCAATCTATGCCATATACATTTCCATCTCTAAAATATGT</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>FJ858801.10n-u</td>
<td>CGTCGGACAACAAATGCAACGTAAGCAACAAAAACATGACAGCGAAGTGGAAAAAGTGCAGCATTAAGA</td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td></td>
<td>-----------------------------------G---TCTATAAAATGTTCACGAGGGGTACCTCCTCTGACATTCAATCAGATGAGGAGTCTTGATCTCGGTGAGGTTGATCTTCGGTGGTGGTGGTGACCACATCTTTTTCCTCTATGCAAAAAAGGATGGGATACTTTGTTAAATTTTGTATTTCAATCTATGCCATATACATTTCCATCTCTAAAATATGT</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>FJ858801.10n-u</td>
<td>CGTCGGACAACAAATGCAACGTAAGCAACAAAAACATGACAGCGAAGTGGAAAAAGTGCAGCATTAAGA</td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td></td>
<td>-----------------------------------G---TCTATAAAATGTTCACGAGGGGTACCTCCTCTGACATTCAATCAGATGAGGAGTCTTGATCTCGGTGAGGTTGATCTTCGGTGGTGGTGGTGACCACATCTTTTTCCTCTATGCAAAAAAGGATGGGATACTTTGTTAAATTTTGTATTTCAATCTATGCCATATACATTTCCATCTCTAAAATATGT</td>
</tr>
<tr>
<td>Accession</td>
<td>Sequence</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>GACAACCTGGGAAAGGCTGCTGCGATGACTGATGTGAGCGACGCTGCTGGCGCCTA</td>
<td></td>
</tr>
<tr>
<td>FJ858801.10n-u</td>
<td>TTGCTCTGCTTGTAGAAAGACCACTGGACAGACGCTGCTGGCTGGCGCCTA</td>
<td></td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td>AGGAAATGGACACAGGCAGGTGTTACCTCTACTCTACCTATTTCCTTGGGACACGACGACTATAATAGT</td>
<td></td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>AAGGGCTCTTCTGCTGTACTTTAGCAACTGCACTGACGTGTCTGACGACGACTATAATAGT</td>
<td></td>
</tr>
<tr>
<td>FJ858801.10n-u</td>
<td>ATGAAAGCAGTGTAGACAGCCAGGCACGAGTCTTCTTACAGAAACATCATATACTGCA</td>
<td></td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td>TTAGAGCTGATTGT-----GCCCCAGAGGACGACGACTATAATAGT</td>
<td></td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>CAGAAACCAGGTTGTTAAT-----GCAAGCCAGGAGCATGTGGCAACTCTGCA</td>
<td></td>
</tr>
<tr>
<td>FJ858801.10n-u</td>
<td>TCGAGCACTAACGACAGAACTGAAGGAAATTTGAAGAAATGCAATGTATGGAAATCAACTGTGCCGTGCAACGC</td>
<td></td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td>GCAAAAAGGGAGTTATGATGCACTGTGGTGGGGGACAGTTATGAGGAACTGCAAGTGGACAAACATGAG</td>
<td></td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>GTATGAAAAGCACTATTGGCTGTCAACTAGAACTGAGACATGATATGACGAGAAGGGAAAATGGGACGACGACTATAATAGT</td>
<td></td>
</tr>
<tr>
<td>FJ858801.10n-u</td>
<td>TCGAGCACTAACGACAGAACTGAAGGAAATTTGAAGAAATGCAATGTATGGAAATCAACTGTGCCGTGCAACGC</td>
<td></td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td>GCAAAAAGGGAGTTATGATGCACTGTGGTGGGGGACAGTTATGAGGAACTGCAAGTGGACAAACATGAG</td>
<td></td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>GACAACCTGGGAAAGGCTGCTGCGATGACTGATGTGAGCGACGCTGCTGGCGCCTA</td>
<td></td>
</tr>
<tr>
<td>FJ858801.10n-u</td>
<td>TTGCTCTGCTTGTAGAAAGACCACTGGACAGACGCTGCTGGCTGGCGCCTA</td>
<td></td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td>AGGAAATGGACACAGGCAGGTGTTACCTCTACTCTACCTATTTCCTTGGGACACGACGACTATAATAGT</td>
<td></td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>AAGGGCTCTTCTGCTGTACTTTAGCAACTGCACTGACGTGTCTGACGACGACTATAATAGT</td>
<td></td>
</tr>
<tr>
<td>FJ858801.10n-u</td>
<td>ATGAAAGCAGTGTAGACAGCCAGGCACGAGTCTTCTTACAGAAACATCATATACTGCA</td>
<td></td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td>TTAGAGCTGATTGT-----GCCCCAGAGGACGACGACTATAATAGT</td>
<td></td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>CAGAAACCAGGTTGTTAAT-----GCAAGCCAGGAGCATGTGGCAACTCTGCA</td>
<td></td>
</tr>
<tr>
<td>FJ858801.10n-u</td>
<td>TCGAGCACTAACGACAGAACTGAAGGAAATTTGAAGAAATGCAATGTATGGAAATCAACTGTGCCGTGCAACGC</td>
<td></td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td>GCAAAAAGGGAGTTATGATGCACTGTGGTGGGGGACAGTTATGAGGAACTGCAAGTGGACAAACATGAG</td>
<td></td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>GACAACCTGGGAAAGGCTGCTGCGATGACTGATGTGAGCGACGCTGCTGGCGCCTA</td>
<td></td>
</tr>
<tr>
<td>FJ858801.10n-u</td>
<td>TTGCTCTGCTTGTAGAAAGACCACTGGACAGACGCTGCTGGCTGGCGCCTA</td>
<td></td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td>AGGAAATGGACACAGGCAGGTGTTACCTCTACTCTACCTATTTCCTTGGGACACGACGACTATAATAGT</td>
<td></td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>AAGGGCTCTTCTGCTGTACTTTAGCAACTGCACTGACGTGTCTGACGACGACTATAATAGT</td>
<td></td>
</tr>
<tr>
<td>FJ858801.10n-u</td>
<td>ATGAAAGCAGTGTAGACAGCCAGGCACGAGTCTTCTTACAGAAACATCATATACTGCA</td>
<td></td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td>TTAGAGCTGATTGT-----GCCCCAGAGGACGACGACTATAATAGT</td>
<td></td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>CAGAAACCAGGTTGTTAAT-----GCAAGCCAGGAGCATGTGGCAACTCTGCA</td>
<td></td>
</tr>
<tr>
<td>FJ858801.10n-u</td>
<td>TCGAGCACTAACGACAGAACTGAAGGAAATTTGAAGAAATGCAATGTATGGAAATCAACTGTGCCGTGCAACGC</td>
<td></td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td>GCAAAAAGGGAGTTATGATGCACTGTGGTGGGGGACAGTTATGAGGAACTGCAAGTGGACAAACATGAG</td>
<td></td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td>Sequence</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td>AY527415.1Foxy</td>
<td>CatCTCCGACTCATCAGCCGAGCTGTCCTGAGCTCTACGAGCGCTTCTCCCTCCCTTTTATG</td>
<td>1123</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>GGTAGTGGCTAGAAAGCCGTAACTTCAGGATCTAGCTGCTCCCAGTCTTCCTTTCTTGCCA</td>
<td>1672</td>
</tr>
<tr>
<td>FJ858801.1On-u</td>
<td>CACACTTCTCATCATCCTGCTCTCAGCCATGGACAGCTCGTTCAGCTTCAGCCCTTT</td>
<td>1732</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>GGAAGATCCAGCGAT---TATCTACAAGCCGAAAAAGGCTCAGCCGCCCTCAGTGCCA</td>
<td>1792</td>
</tr>
<tr>
<td>FJ858801.1On-u</td>
<td>GTGATGGCTAGAAAGCCGTAACTTCAGGATCTAGCTGCTCCCAGTCTTCCTTTCTTGCCA</td>
<td>1792</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>AAACGGGTAGCTATCGACTCAGCCCGAGGGGCTGGAGATGGTGACGTGGTGGT</td>
<td>1851</td>
</tr>
<tr>
<td>FJ858801.1On-u</td>
<td>AAACGGGTAGCTATCGACTCAGCCCGAGGGGCTGGAGATGGTGACGTGGTGGT</td>
<td>1851</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>CTCGCCGCTCAACATCATTGATGCCGCCCTTGTGAGATGGTATACTGGAGCTGTCGTTGT</td>
<td>1909</td>
</tr>
<tr>
<td>FJ858801.1On-u</td>
<td>CTGAGACCTGTAAATGATCCAGGCATTAGGCGGGCGAGGGATGCGTGACTCGGTTGTGGT</td>
<td>1969</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>CATCGATAATGAGCTCCTCAACGGAATCCGAGCTACTCATATCAGATTGGCCAGGTATGG</td>
<td>2029</td>
</tr>
<tr>
<td>FJ858801.1On-u</td>
<td>TATATACAAATGGTCTGTTACCTGCACTTACCTTTGGCATGGCCGATGATCTCAAAGCGG</td>
<td>2089</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>TCGAGGAGTCGCACTCAGCCCTCTGGCTGAGCTTTGAGGTGCAGCAGGATGTGAGC</td>
<td>2149</td>
</tr>
<tr>
<td>FJ858801.1On-u</td>
<td>TCGTAGACA---------ATGGAGCAATCCTCATTGAGTGATCCGAGACTGATGG</td>
<td>2200</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>TCGGCCGCTCAACATCATTGATGCCGCCCTTGTGAGATGGTATACTGGAGCTGTCGTTGT</td>
<td>2260</td>
</tr>
<tr>
<td>FJ858801.1On-u</td>
<td>GATGAGTATGCGAGGAGAATAGTGTGTTTCTCAAGACTTCTCTACTGAGGTGCAGGATGG</td>
<td>2320</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>CATTGACAACAAGTGTCACTCGGAGCTCGATAGTGGTAGGCATTATGTGGTGATGATGTT</td>
<td>2380</td>
</tr>
<tr>
<td>FJ858801.1On-u</td>
<td>GAGTATGGCTGCTCGAATTCGTGGGCTGCTCTGCTGACTTGATGTTGGTGCTG</td>
<td>2440</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>TCCGGCTGAGGACACTTGCTTCTAAGGACACTGACTGATGGTATTATGATGG</td>
<td>2500</td>
</tr>
<tr>
<td>FJ858801.1On-u</td>
<td>CATTGACAACAAGTGTCACTCGGAGCTCGATAGTGGTAGGCATTATGTGGTGATGATG</td>
<td>2560</td>
</tr>
<tr>
<td>AM983455.1Achr</td>
<td>CAGTATGGCTGCTCGAATTCGTGGGCTGCTCTGCTGACTTGATGTTGGTGCTG</td>
<td>2620</td>
</tr>
<tr>
<td>FJ858801.1On-u</td>
<td>CAGTATGGCTGCTCGAATTCGTGGGCTGCTCTGCTGACTTGATGTTGGTGCTG</td>
<td>2680</td>
</tr>
</tbody>
</table>

346
Appendix 5.2 Alignment of MAT1-2-1 sequences collected from GenBank for primer design.

Sequence outside of primer design region was truncated for simplicity.
AB080673.2Mgri  G----------------------------------------------- 459
AY357890.1Gcin  ATGTACTCACAACCCATGTATTCCGATGTCAAGCGAATACGGATCTTTTTCTGTAATTAT  892

AB080673.2Mgri  -----ACAGTGCAAGTCGATGTCAGC--------------------------------- 479
AY357890.1Gcin  AATCTGCAAGCCAGTAAATATCATCTTCATGGACCATCTATTCCAGCCCGATACCTAGA  952
  **   ****  ** *** *

354
Appendix 5.3 Alignment of SLA2 coding sequences. Note that trailing sequence has been truncated.
Appendix 5.5 Alignment of APN2 coding sequences

<table>
<thead>
<tr>
<th>Sequence 1</th>
<th>Sequence 2</th>
<th>Sequence 3</th>
<th>Sequence 4</th>
<th>Sequence 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>g10805_mm10095_APN2</td>
<td>g6680_mm10106_APN2</td>
<td>g6890_mb07020_APN2</td>
<td>g9067.t1_mm99049_apn2</td>
<td>g816.t1_mm12262_apn2</td>
</tr>
<tr>
<td>g10269.t1_mm11037_apn2</td>
<td></td>
<td></td>
<td>g10269.t1_mm11037_apn2</td>
<td></td>
</tr>
<tr>
<td>g10805_mm10095_APN2</td>
<td>g6680_mm10106_APN2</td>
<td>g6890_mb07020_APN2</td>
<td>g9067.t1_mm99049_apn2</td>
<td>g816.t1_mm12262_apn2</td>
</tr>
<tr>
<td>g10269.t1_mm11037_apn2</td>
<td></td>
<td></td>
<td>g10269.t1_mm11037_apn2</td>
<td></td>
</tr>
<tr>
<td>g10805_mm10095_APN2</td>
<td>g6680_mm10106_APN2</td>
<td>g6890_mb07020_APN2</td>
<td>g9067.t1_mm99049_apn2</td>
<td>g816.t1_mm12262_apn2</td>
</tr>
<tr>
<td>g10269.t1_mm11037_apn2</td>
<td></td>
<td></td>
<td>g10269.t1_mm11037_apn2</td>
<td></td>
</tr>
<tr>
<td>g10805_mm10095_APN2</td>
<td>g6680_mm10106_APN2</td>
<td>g6890_mb07020_APN2</td>
<td>g9067.t1_mm99049_apn2</td>
<td>g816.t1_mm12262_apn2</td>
</tr>
<tr>
<td>g10269.t1_mm11037_apn2</td>
<td></td>
<td></td>
<td>g10269.t1_mm11037_apn2</td>
<td></td>
</tr>
</tbody>
</table>

---

**Alignment of APN2 coding sequences**

---

**366**
<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>g10805_mm10095_APN2</td>
<td>tggagcagcgactggactagcaagagc</td>
</tr>
<tr>
<td>g6680_mm10106_APN2</td>
<td>tggagcagcgactggactagcaagagc</td>
</tr>
<tr>
<td>g6890_mb07020_APN2</td>
<td>tggagcagtgactggactagcaagagc</td>
</tr>
<tr>
<td>g9067.t1_mm99049_apn2</td>
<td>tggagcagcgactggactagcaagagc</td>
</tr>
<tr>
<td>g816.t1_mm12262_apn2</td>
<td>tggagcagtgactggactagcaagagc</td>
</tr>
<tr>
<td>g10269.t1_mm11037_apn2</td>
<td>tggagcagtgactggactagcaagagc</td>
</tr>
</tbody>
</table>

***** ** *******

* ***** ***
## Appendix 5.6 Alignment of *M. majus* 99049 MAT-region coding sequences with those of *M. nivale* 11037, with primer loci indicated for SLA2

<table>
<thead>
<tr>
<th>Sequence Name</th>
<th>Alignment</th>
<th>Primer Loci</th>
</tr>
</thead>
<tbody>
<tr>
<td>g814.t1_mm12262_sla2</td>
<td>ATGGCGTCCGGCCAGTCTGACAAAGTCGGAGCCGAGGC 50</td>
<td></td>
</tr>
<tr>
<td>g9065.t1_mm99049_sla2</td>
<td>ATGGCGCTCCGGCCAGTCTGACAAAGTCGGAGCCGAGGC 50</td>
<td></td>
</tr>
<tr>
<td>g10271.t1_mm11037_sla2</td>
<td>ATGGCGTCCGGCCAGTCTGACAAAGTCGGAGCCGAGGC 50</td>
<td></td>
</tr>
<tr>
<td>g814.t1_mm12262_sla2</td>
<td>CATCAACATCAAAAAGGCTACGAGTCCCGAGGAGTCGGCCCAGCAGCGCA 100</td>
<td></td>
</tr>
<tr>
<td>g9065.t1_mm99049_sla2</td>
<td>CATTAACATAAAAAAGGCCACAAGTCCCGAGGAGTCGGCCCAGCAGCGCA 100</td>
<td></td>
</tr>
<tr>
<td>g10271.t1_mm11037_sla2</td>
<td>CATCAATATCAAAAAAGGCCACAAGTCCCGAGGAGTCGGCCCAGCAGCGCA 100</td>
<td></td>
</tr>
<tr>
<td>g814.t1_mm12262_sla2</td>
<td>CATCAACATCAAAAAGGCTACGAGTCCCGAGGAGTCGGCCCAGCAGCGCA 100</td>
<td></td>
</tr>
<tr>
<td>g9065.t1_mm99049_sla2</td>
<td>CATTAACATAAAAAAGGCCACAAGTCCCGAGGAGTCGGCCCAGCAGCGCA 100</td>
<td></td>
</tr>
<tr>
<td>g10271.t1_mm11037_sla2</td>
<td>CATCAATATCAAAAAAGGCCACAAGTCCCGAGGAGTCGGCCCAGCAGCGCA 100</td>
<td></td>
</tr>
<tr>
<td>g814.t1_mm12262_sla2</td>
<td>CATCAACATCAAAAAGGCTACGAGTCCCGAGGAGTCGGCCCAGCAGCGCA 100</td>
<td></td>
</tr>
<tr>
<td>g9065.t1_mm99049_sla2</td>
<td>CATTAACATAAAAAAGGCCACAAGTCCCGAGGAGTCGGCCCAGCAGCGCA 100</td>
<td></td>
</tr>
<tr>
<td>g10271.t1_mm11037_sla2</td>
<td>CATCAATATCAAAAAAGGCCACAAGTCCCGAGGAGTCGGCCCAGCAGCGCA 100</td>
<td></td>
</tr>
<tr>
<td>g814.t1_mm12262_sla2</td>
<td>CATCAACATCAAAAAGGCTACGAGTCCCGAGGAGTCGGCCCAGCAGCGCA 100</td>
<td></td>
</tr>
<tr>
<td>g9065.t1_mm99049_sla2</td>
<td>CATTAACATAAAAAAGGCCACAAGTCCCGAGGAGTCGGCCCAGCAGCGCA 100</td>
<td></td>
</tr>
<tr>
<td>g10271.t1_mm11037_sla2</td>
<td>CATCAATATCAAAAAAGGCCACAAGTCCCGAGGAGTCGGCCCAGCAGCGCA 100</td>
<td></td>
</tr>
<tr>
<td>g814.t1_mm12262_sla2</td>
<td>AGCACATACAAGGCGCTCATCACCATTCACAAGTCCCTCCCAG 150</td>
<td></td>
</tr>
<tr>
<td>g9065.t1_mm99049_sla2</td>
<td>AGCACATACAAGGCGCTCATCACCATTCACAAGTCCCTCCCAG 150</td>
<td></td>
</tr>
<tr>
<td>g10271.t1_mm11037_sla2</td>
<td>AGCATGTCCGCAGCTGCATCGTCTACACATGGGACCACAAGTCCTCCCAG 150</td>
<td></td>
</tr>
<tr>
<td>g814.t1_mm12262_sla2</td>
<td>TCCTTCTGGGCTGGGCTCAAGGTGCAGCCTATCCTCGCCGACGAGGTCCA 200</td>
<td></td>
</tr>
<tr>
<td>g9065.t1_mm99049_sla2</td>
<td>TCCTTCTGGGCTGGGCTCAAGGTGCAGCCTATCCTCGCCGACGAGGTCCA 200</td>
<td></td>
</tr>
<tr>
<td>g10271.t1_mm11037_sla2</td>
<td>TCCTTCTGGGCTGGGCTCAAGGTGCAGCCTATCCTCGCCGACGAGGTCCA 200</td>
<td></td>
</tr>
<tr>
<td>g814.t1_mm12262_sla2</td>
<td>GACATACAAGGCGCTCATCACCATTCACAAGGTCCTCCAAGAGGGCCACC 250</td>
<td></td>
</tr>
<tr>
<td>g9065.t1_mm99049_sla2</td>
<td>GACATACAAGGCGCTCATCACCATTCACAAGGTCCTCCAAGAGGGCCACC 250</td>
<td></td>
</tr>
<tr>
<td>g10271.t1_mm11037_sla2</td>
<td>GACATACAAGGCGCTCATCACCATTCACAAGGTCCTCCAAGAGGGCCACC 250</td>
<td></td>
</tr>
<tr>
<td>g814.t1_mm12262_sla2</td>
<td>CGCAAACCCTCAGAGAGGCGATGGCCAACCCGAGCTGGATCGACAGCCTG 300</td>
<td></td>
</tr>
<tr>
<td>g9065.t1_mm99049_sla2</td>
<td>CGCAAACCCTCAGAGAGGCGATGGCCAACCCGAGCTGGATCGACAGCCTG 300</td>
<td></td>
</tr>
<tr>
<td>g10271.t1_mm11037_sla2</td>
<td>CGCAAACCCTCAGAGAGGCGATGGCCAACCCGAGCTGGATCGACAGCCTG 300</td>
<td></td>
</tr>
<tr>
<td>g814.t1_mm12262_sla2</td>
<td>AATAGAGGCATGAGCGGCGAGGGTATGCGTGGATATGCCCCTCTCATTCG 350</td>
<td></td>
</tr>
<tr>
<td>g9065.t1_mm99049_sla2</td>
<td>AATAGAGGCATGAGCGGCGAGGGTATGCGTGGATATGCCCCTCTCATTCG 350</td>
<td></td>
</tr>
<tr>
<td>g10271.t1_mm11037_sla2</td>
<td>AATAGAGGCATGAGCGGCGAGGGTATGCGTGGATATGCCCCTCTCATTCG 350</td>
<td></td>
</tr>
<tr>
<td>g814.t1_mm12262_sla2</td>
<td>CTGAGTTTAACGGCACCTTCGAGTACGAAGAGTACGTCTCGCTCAAGGCC 450</td>
<td></td>
</tr>
<tr>
<td>g9065.t1_mm99049_sla2</td>
<td>CTGAGTTTAACGGCACCTTCGAGTACGAAGAGTACGTCTCGCTCAAGGCC 450</td>
<td></td>
</tr>
<tr>
<td>g10271.t1_mm11037_sla2</td>
<td>CTGAGTTTAACGGCACCTTCGAGTACGAAGAGTACGTCTCGCTCAAGGCC 450</td>
<td></td>
</tr>
<tr>
<td>g814.t1_mm12262_sla2</td>
<td>ATGTCGGCAACAAGATCGACCAGTTCCAGAAGCTCATCTTCTCACACTTCCGCA 550</td>
<td></td>
</tr>
<tr>
<td>g9065.t1_mm99049_sla2</td>
<td>ATGTCGGCAACAAGATCGACCAGTTCCAGAAGCTCATCTTCTCACACTTCCGCA 550</td>
<td></td>
</tr>
<tr>
<td>g10271.t1_mm11037_sla2</td>
<td>ATGTCGGCAACAAGATCGACCAGTTCCAGAAGCTCATCTTCTCACACTTCCGCA 550</td>
<td></td>
</tr>
</tbody>
</table>

**Mn_SLA2_357F:**

<table>
<thead>
<tr>
<th>Sequence Name</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>g814.t1_mm12262_sla2</td>
<td>ATGTCGGCAACAAGATCGACCAGTTCCAGAAGCTCATCTTCTCACACTTCCGCA 550</td>
</tr>
<tr>
<td>g9065.t1_mm99049_sla2</td>
<td>ATGTCGGCAACAAGATCGACCAGTTCCAGAAGCTCATCTTCTCACACTTCCGCA 550</td>
</tr>
<tr>
<td>g10271.t1_mm11037_sla2</td>
<td>ATGTCGGCAACAAGATCGACCAGTTCCAGAAGCTCATCTTCTCACACTTCCGCA 550</td>
</tr>
</tbody>
</table>

**Mic_SLA2_92F:**

<table>
<thead>
<tr>
<th>Sequence Name</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>g814.t1_mm12262_sla2</td>
<td>ATGTCGGCAACAAGATCGACCAGTTCCAGAAGCTCATCTTCTCACACTTCCGCA 550</td>
</tr>
<tr>
<td>g9065.t1_mm99049_sla2</td>
<td>ATGTCGGCAACAAGATCGACCAGTTCCAGAAGCTCATCTTCTCACACTTCCGCA 550</td>
</tr>
<tr>
<td>g10271.t1_mm11037_sla2</td>
<td>ATGTCGGCAACAAGATCGACCAGTTCCAGAAGCTCATCTTCTCACACTTCCGCA 550</td>
</tr>
<tr>
<td>Gene1</td>
<td>Gene2</td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>g9065.t1_mm99049_sla2</td>
<td>GTGGCTCTGTCGACCATGCTTTATAACAGCACTGACAGACTCGAGACATGACG 2000</td>
</tr>
<tr>
<td>g10271.t1_mm11037_sla2</td>
<td>GTGGCTCTGTCGACCATGCTTTATAACAGCACTGACAGACTCGAGACATGACG 2000</td>
</tr>
<tr>
<td>g814.t1_mm112262_sla2</td>
<td>TGGTATTTTAACAGACGCTTGGACCATGCTTTATAACAGCACTGACAGACTCGAGACATGACG 2005</td>
</tr>
<tr>
<td>g9065.t1_mm99049_sla2</td>
<td>TGGTATTTTAACAGACGCTTGGACCATGCTTTATAACAGCACTGACAGACTCGAGACATGACG 2005</td>
</tr>
<tr>
<td>g10271.t1_mm11037_sla2</td>
<td>TGGTATTTTAACAGACGCTTGGACCATGCTTTATAACAGCACTGACAGACTCGAGACATGACG 2005</td>
</tr>
</tbody>
</table>

** Additional sequences omitted for brevity. **
### Appendix 5.7 Alignment of M. majus 99049 MAT-region coding sequences with those of *M. nivale* 11037, with primer loci indicated for MAT1-2-1

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Primer Loci</th>
<th>Sequence Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>g815.t1_mn12262_mat2</td>
<td>CTCAATCTTGTGACCGGCTGCTTCTGCTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
<tr>
<td>g9066.t1_mn99049_mat2</td>
<td>CTCACAGTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
<tr>
<td>g10270.t1_mn11037_mat2</td>
<td>CTCACAGTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
<tr>
<td>g815.t1_mn12262_mat2</td>
<td>CTCAATCTTGTGACCGGCTGCTTCTGCTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
<tr>
<td>g9066.t1_mn99049_mat2</td>
<td>CTCACAGTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
<tr>
<td>g10270.t1_mn11037_mat2</td>
<td>CTCACAGTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
<tr>
<td>g815.t1_mn12262_mat2</td>
<td>CTCAATCTTGTGACCGGCTGCTTCTGCTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
<tr>
<td>g9066.t1_mn99049_mat2</td>
<td>CTCACAGTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
<tr>
<td>g10270.t1_mn11037_mat2</td>
<td>CTCACAGTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
<tr>
<td>g815.t1_mn12262_mat2</td>
<td>CTCAATCTTGTGACCGGCTGCTTCTGCTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
<tr>
<td>g9066.t1_mn99049_mat2</td>
<td>CTCACAGTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
<tr>
<td>g10270.t1_mn11037_mat2</td>
<td>CTCACAGTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
<tr>
<td>g815.t1_mn12262_mat2</td>
<td>CTCAATCTTGTGACCGGCTGCTTCTGCTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
<tr>
<td>g9066.t1_mn99049_mat2</td>
<td>CTCACAGTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
<tr>
<td>g10270.t1_mn11037_mat2</td>
<td>CTCACAGTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
<tr>
<td>g815.t1_mn12262_mat2</td>
<td>CTCAATCTTGTGACCGGCTGCTTCTGCTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
<tr>
<td>g9066.t1_mn99049_mat2</td>
<td>CTCACAGTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
<tr>
<td>g10270.t1_mn11037_mat2</td>
<td>CTCACAGTGACAGTGGTACCCAGAACAAG</td>
<td>ATCGAGGTCAGCTGCTGACAGTGGTACCCAGAACAAG</td>
</tr>
</tbody>
</table>

*Note: The sequences are presented in the 5' to 3' direction and the primer loci are indicated for MAT1-2-1.*
<table>
<thead>
<tr>
<th>Code</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>g815.t1_mn12262_mat2</td>
<td>AACTAC 1941</td>
</tr>
<tr>
<td>g9066.t1_mm99049_mat2</td>
<td>AACTAC 1860</td>
</tr>
<tr>
<td>g10270.t1_mn11037_mat2</td>
<td>AACTAC 1671</td>
</tr>
<tr>
<td></td>
<td>*****</td>
</tr>
</tbody>
</table>
Appendix 5.8 Alignment of M. majus 99049 MAT-region coding sequences with those of M. nivale 11037, with primer loci indicated for APN2
<table>
<thead>
<tr>
<th>Accession</th>
<th>Sequence</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>g016.t1_mn12262_apn2</td>
<td>CTAGCAAGAC</td>
<td>1866</td>
</tr>
<tr>
<td>g9067.t1_mm99049_apn2</td>
<td>CTAGCAAGAC</td>
<td>1824</td>
</tr>
<tr>
<td>g10269.t1_mn11037_apn2</td>
<td>CTAGCAAAAGC</td>
<td>1962</td>
</tr>
</tbody>
</table>

******* ***
### Appendix 5.9 Observed ascospore dimensions

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Length (µm)</th>
<th>Width (µm)</th>
<th>Q*</th>
<th>Number of Septa</th>
</tr>
</thead>
<tbody>
<tr>
<td>99049</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16.8</td>
<td>4.8</td>
<td>3.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>24.0</td>
<td>4.8</td>
<td>5.0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>14.4</td>
<td>4.8</td>
<td>3.0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>14.4</td>
<td>3.6</td>
<td>4.0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>14.4</td>
<td>4.8</td>
<td>3.0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>21.6</td>
<td>4.8</td>
<td>4.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>19.2</td>
<td>4.8</td>
<td>4.0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>16.8</td>
<td>4.8</td>
<td>3.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>16.8</td>
<td>4.8</td>
<td>3.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>16.8</td>
<td>4.8</td>
<td>3.0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>14.4</td>
<td>4.8</td>
<td>3.0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>19.2</td>
<td>6.0</td>
<td>3.2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>19.2</td>
<td>4.8</td>
<td>4.0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>16.8</td>
<td>4.8</td>
<td>3.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>12.0</td>
<td>4.8</td>
<td>2.5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>19.2</td>
<td>6.0</td>
<td>3.2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>16.8</td>
<td>4.8</td>
<td>3.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>16.8</td>
<td>3.6</td>
<td>4.7</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>19.2</td>
<td>2.4</td>
<td>8.0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>16.8</td>
<td>3.6</td>
<td>4.7</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>19.2</td>
<td>3.6</td>
<td>5.3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>14.4</td>
<td>3.6</td>
<td>4.0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>12.0</td>
<td>3.6</td>
<td>3.3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>14.4</td>
<td>4.8</td>
<td>3.0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>16.8</td>
<td>4.8</td>
<td>3.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>12166</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16.8</td>
<td>4.8</td>
<td>3.5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>16.8</td>
<td>4.8</td>
<td>3.5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>12.0</td>
<td>4.8</td>
<td>2.5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>14.4</td>
<td>4.8</td>
<td>3.0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>16.8</td>
<td>4.8</td>
<td>3.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>14.4</td>
<td>4.8</td>
<td>3.0</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

* Q = length / width
Appendix 5.10 Supplemental Results: Test of published mating type primers and redesigned universal primers

The MAT1_3149F and MAT1_3644R primer set designed to amplify MAT1-1-1 yielded a single band of the predicted size (300 bp) from two of the four isolates it was initially tested with. When one of these bands (from *M. nivale* isolate 10083) was sequenced in the forward direction, the resulting amplicon was queried against the GenBank database (BLASTx and BLASTn) to assess its similarity to other MAT1-1-1 sequences in the database. The top match in both cases was the protein eukaryotic release factor 1 which has not been described as having a role in sexual reproduction. The MAT2_3515F and MAT2_4618R primer set failed to yield a single band with the isolates with which they were tested.

Two sets of redesigned primers, MAT1_3132F and MAT1_3303R, and MAT2_1404F and MAT2_1811R were also designed and tested, but failed to yield bands of the predicted size from *M. nivale* or *M. majus*. Another set of mating type primers, MAT1_86F and MAT1_891R, and MAT2_488F and MAT2_650R were also designed and tested. The MAT1_86F and MAT1_891R primer set produced strong bands of the predicted size (805 bp), as well as larger bands of approximately 900 bp in two of the five isolates it was tested with. After optimization, single bands of approximately 1,500 bp and 600 bp were produced for the *M. majus* isolate 10096 and the *M. nivale* isolate 10102, respectively. Both were sequenced in the forward direction, but the resulting sequences were of poor quality, suggesting that multiple bands of approximately the same size had been produced by the primer.

The MAT1-2-1 primers MAT2_488F and MAT2_650R produced a single band of approximately 600 bp in all five of the isolates it was tested with. Although this band was larger than the predicted amplicon, it was sent for sequencing in the forward direction. The resulting
sequence had no matches in the GenBank database using the BLASTn algorithm, and the BLASTx algorithm yielded only a weak set of matches. The closest match in the BLASTx search (minimum e-value 0.26 across 24% of the query) was a hypothetical protein from Magnaporthe grisea, which in turn shared a high homology with a GATA transcription factor.
Appendix 5.11 Supplemental results: Identification of flanking genes in May 2011 99049 assembly

In tBLASTn searches against the original assembly of the *M. majus* genome, the putative match to SLA2 displayed a 68.9% identity and an e-value of 0.0 with the query. For MAT1-2-1, the putative match displayed a 31.5% identity and an e-value of 4e-06 with the query sequence from *Stemphylium majusculum*. For APN2, the putative match displayed a 53.3% identity and an e-value of 7e-119.

Primers were designed near the 3' end of the putative APN2 sequence and the 5' end of the putative SLA2 sequence, with the goal of amplifying the region spanning between these genes, which was hypothesized to include the putative MAT1-2-1 gene identified as described above. These SLA2 and APN2 primers, Mn_APN2_838F and Mn_SLA2_23R, were tested with four *M. majus* isolates, including the sequenced isolate 99049. Despite manipulating the PCR conditions as described, these primers failed to amplify a fragment of the predicted size (between 6-8 kb, based on the size of this region in other species).


Chapter 6 Infection Process of *Microdochium majus* and *M. nivale*

6.1 Introduction

6.1.1 The disease cycle

The interaction between a given fungal pathogen and its host involves a series of distinct events known as the disease cycle (Agrios 2005). These events (Figure 6.1) include the dissemination of the infectious material (inoculum) to the potential host plant, the attachment to and penetration of the plant tissue, the infection and invasion of the host plant cells, and the growth of the pathogen that ultimately produces a fresh source of inoculum for subsequent infections. Depending on the identity of the pathogen, the inoculum may consist of hyphal fragments, conidia, sexual spores, or resting structures (e.g. sclerotia) (Agrios 2005), and the inoculum may be transferred to the host plant by diverse mechanisms including wind, water, insects, animals, or agricultural practices.

The series of physical processes that occur during the invasion of a plant by a pathogen are generally referred to as the infection process for a particular pathosystem. To invade the host plant, the pathogen may take advantage of existing openings, such as wounds created by insects or by human activity, or natural openings such as stomata. Alternatively, the fungus may directly penetrate into cells. Although it is possible for actively growing hyphae to infiltrate the plant tissue by simply growing through an existing opening, in many fungi, the infection process begins when conidia come into contact with the surface of a susceptible host plant. In some pathogens, following germination of the conidium, the tip of the emergent germ tube may differentiate into a specialized structure called an appressorium (Howard 1996), which provides the attachment necessary to drive a penetration peg either directly through the plant cuticle or
through an existing opening. When appressoria are not formed, penetration may be more heavily reliant on the enzymatic degradation of the plant tissue (Mendgen et al. 1996).

Once inside the plant, the fungus may infiltrate neighbouring cells, causing structural damage and robbing the host plant of resources. While some plant pathogens are biotrophs, requiring a living host (Agrios 2005), saprotrophs are capable of growing on dead tissue. Some pathogens exhibit a mixture of these behaviours by beginning their infection as biotrophs, but switching to a saprotrophic lifestyle after the surrounding plant cells die or are killed (Agrios 2005).

To absorb nutrients from their host plants, biotrophs may develop specialized structures called haustoria which penetrate the plant cell wall, but not its membrane (Mendgen et al. 2000). The developing haustorium expands in shape, often developing “finger-like” projections which maximize the surface area between the fungus and the plant cell membrane for the purpose of nutrient acquisition from the plant cell (Heath and Skalamera 1997). Although several genera within the Xylariales include phytopathogenic and / or endophytic members (Brunner and Petrini 1992; Davis et al. 2003), few reports are available describing the infection processes of these species.

6.1.2 The infection processes of *M. nivale* and *M. majus*

While conidia are important in the dissemination of many plant diseases, including other diseases of graminoid plants (e.g. Colletotrichum graminicola, the causal agent of anthracnose diseases of grasses (Khan and Hsiang 2006)), the conidia of Microdochium nivale sensu lato are not reported to initiate disease symptoms (Pronczuk and Messyasz 1991; Tronsmo et al. 2001). Outbreaks of this pathogen appear to be mediated by the spread of soil-borne hyphae, as conidial germination and subsequent germ tube formation appears to be a prohibitively slow and non-
infective process (Pronczuk and Messyasz 1991). This unusual observation may be explained at least in part by the reduced levels of competition faced by *M. nivale* and *M. majus* during the cool temperatures in which they cause disease (Tronsmo et al. 2001). It may also be possible that the conidia are not directly infective, and may instead germinate in the soil; these species may then survive as saprotrophs, only behaving as pathogens when under favourable environmental conditions (Tronsmo et al. 2001). Alternatively, conidia may play a role in sexual reproduction by transferring genetic information between strains, performing a spermatizing function (Webster and Weber 2007).

The infection process of *M. nivale* sensu lato has been studied on three turfgrass species (Dahl 1934). On all host species described, fungal hyphae penetrated through stomata, rather than directly punching through the foliar surface (Dahl 1934). Penetration occurred within 3 days post-inoculation when hyphal inoculum was used. A study of the infection process of *M. nivale* on the cereal triticale (*x Triticosecale*, a cross of wheat, *Triticum vulgare*, with rye, *Secale cereale*) was recently reported (Dubas et al. 2011). By using fluorescent dyes to track the growth of the fungus following hyphal inoculation, these researchers found that hyphae grew directly into the stomata of the plants, and formed haustorium-like structures within the plant cells. Haustoria are a typical feature of biotrophic fungi (Catanzariti 2011), and have not been previously reported for *M. nivale*. However, both studies on triticale were performed with the same strain of *M. nivale* that was originally isolated from rye in 2001, and although the authors state that this particular strain was highly virulent, it is unclear whether it was truly *M. nivale*, or actually *M. majus*.

On both rye and triticale, researchers have reported the presence of haustoria-like structures in infected tissues (Dubas et al. 2011; Zur et al. 2011). The presence of haustoria implies that
these pathogens are engaged in biotrophy, absorbing nutrition from the living host plant. *Microdochium nivale* sensu lato has previously been described as saprotrophic, as it is capable of surviving on dead plant material without a living host for up to one year (Tronsmo et al. 2001). It is possible that *M. nivale* and *M. majus* are hemibiotrophs, exhibiting a short period of biotrophy on the living host before the host tissue is killed by degradative enzymes (Agrios 2005), but this needs to be confirmed by more extensive microscopic and physiological work.

### 6.1.3 Host specificity and infection success

*Microdochium nivale* and *M. majus* are known to possess host preferences, in that *M. majus* has only been isolated from cereals whereas *M. nivale* colonizes both cereals and grasses (Hofgaard et al. 2006; Ioos et al. 2004; Mahuku et al. 1998; Simpson et al. 2000; Smith 1983). The ability of each species to preferentially colonize different hosts may be related to their differing abilities to overcome host defences or due to their ability to avoid initiating or inducing different levels or types of defensive responses in different hosts (Simpson et al. 2000). Similarly, it may be possible that the physical processes of stomatal infiltration may differ between *M. nivale* and *M. majus*, either in general or when in contact with different hosts.

### 6.1.4 Objectives

The purpose of this research was to determine whether the timing and mechanism of the infection processes of *M. nivale* and *M. majus* differ on different host plants, and to determine whether within-species variation might be observed for *M. nivale* isolates initially collected from either wheat (*Triticum aestivum*) or from a turfgrass species. Additionally, the infectivity of conidia and hyphal fragments was also investigated. Finally, the claim that specialized structures
such as haustoria are produced during infection was investigated on detached leaves of wheat and Kentucky bluegrass (*Poa pratensis*) by searching for the presence of these structures on inoculated leaves.

### 6.2 Materials and Methods

#### 6.2.1 Plant culture

Seeds of *Triticum aestivum* (red hard winter wheat) or *Poa pratensis* (Kentucky bluegrass) cultivar “common #1” were sterilized and pre-germinated before planting. The seeds were rinsed in sterile distilled H$_2$O (SDW) for 30 s, then in 70% EtOH for 15 s before being sterilized in 1% NaOCl and finally rinsed in SDW. Seeds of wheat were held in the sterilization solution for 1 min, while Kentucky bluegrass seeds were sterilized for 30 s. The seeds were then placed in Petri plates containing two sheets of autoclaved 7 cm-diameter filter paper (Whatman qualitative circles, grade 2; Fisher Scientific) moistened with SDW. After 48 hours, germinated seeds were planted in 500 mL Mason jars (Bernardin, Richmond Hill, ON, Canada) containing 50 mL (300 g) of rooting mix autoclaved twice (20 minute cycle at 121°C) at 24 hr intervals. The rooting mix contained 80% sand and 20% peat. Approximately 100 seeds of Kentucky bluegrass or 30 seeds of wheat were planted in each jar. The plants were maintained at 23°C for three weeks under high humidity and under constant fluorescent lights emitting 50 μmol/m$^2$/s, and were watered with SDW as necessary.

#### 6.2.2 Inoculum preparation

To induce conidiation for spore inoculation, agar plugs were cut from the actively-growing margin of a colony, and one was placed in the centre of a fresh PDA plate. The plates were
incubated at room temperature in the dark for 48 hours before being exposed to UV light at room temperature. For isolates that readily produced conidia, conidiation was induced within 48 hours of exposure to UV light. Some isolates failed to produce conidia, and hence different isolates were used in the different experiments (Table 6.1). Spore suspensions were prepared by pipetting 1.5 mL of SDW into each plate of an actively sporulating colony, gently scratching with a sterile glass rod, then filtering this suspension through two layers of autoclaved muslin into a 1.5 mL plastic tube. The spore concentration was determined using a haemocytometer, and the final concentration was adjusted to $10^6$ spores / mL using SDW.

For experiment A, an auger was used to cut agar plugs 5 mm in diameter from a two-week old colony of the isolate of interest growing on PDA. Hyphal inoculum was prepared by pouring 3 mL of SDW onto the surface of a non-sporulating two-week-old colony of the isolate of interest growing on PDA. The plates were scraped with a sterile glass rod, and the resulting hyphal suspension was diluted 1:10 with autoclaved water.

6.2.3 Inoculation

Five separate experiments, denoted as A-E, were performed. A general summary of the types of inoculum and the number of isolates included in each experiment is available in Table 6.1, and a list of isolates used in each experiment is provided in Table 6.2. Both host plant species were inoculated with every fungal isolate used in each experiment. All plates for each host plant by isolate combination were prepared in triplicate. Non-inoculated controls were prepared using either agar plugs from fresh (non-inoculated) PDA plates (experiments A and B) or by spraying the leaves with SDW (experiments C-E). To ensure that they were viable, the inocula used in these experiments were also applied to fresh PDA plates and the growth was monitored.
Leaf blades were collected from the wheat and Kentucky bluegrass grown as described (Section 6.2.1), and were cut into sections approximately 1 cm in length using a flame-sterilized scalpel. The leaf sections were placed in Petri plates containing two sheets of autoclaved 7 cm-diameter filter paper (Whatman qualitative circles, grade 2; Fisher Scientific) moistened with SDW. Approximately 20 sections were placed in each plate, and each plate received a single type of inoculum Table 6.1). For hyphal inoculation in experiments A and B, a plug of PDA 3mm in diameter cut from the actively-growing region of a colony was placed face-down onto each leaf segment. The plug was removed after 24 hours. For experiments C-E, leaf blade sections were inoculated with approximately 200 µL of the hyphal or conidial suspension using a hand pump sprayer.

Immediately following inoculation, the plates containing the inoculated leaf sections were placed in a plastic bin and were incubated under the conditions described in Table 6.1. For the "full light" condition, plates were placed in a transparent plastic box and incubated in the presence of a fluorescent light emitting 50 µmol/m²/s for 24 h each day. For the "darkness" condition, plates were placed in an opaque plastic box. To maintain high humidity, paper towels saturated with sterile water were placed along the bottom of the box and re-moistened as necessary.

6.2.4 Sample collection, staining, and scoring

At each collection time point (Table 6.3), one leaf segment was removed at random from each of the plates. All three segments from a single treatment group were placed in 1.0 mL of acetic alcohol (1:3 glacial acetic acid: 95% ethanol) for 24 h to remove chlorophyll from the leaves. After 24 h the acetic alcohol was carefully poured off and fresh acetic alcohol was added.
After a second 24 h period, the acetic alcohol was removed and the leaf sections were placed onto a glass slide containing two drops (approximately 100 μL) of 0.05% trypan blue (w/v) in lactophenol (20% phenol, 20% lactic acid, 40% glycerine, and 20% water). The leaf sections were stained for 48 h before being soaked in lactophenol for 12-24 h to remove excess stain. Finally, the segments were mounted in two drops of lactophenol and were covered by a glass coverslip. A layer of transparent nail polish was painted around the exterior of the coverslip to prevent movement or drying of the mounted material. Due to the destructive nature of this sampling protocol, different leaf segments were examined at each time point.

The slides were examined at 100x and 400x magnification with a Nikon Labophot microscope, and representative photographs were taken using either a Nikon CoolPix800 digital camera (experiments A and B) or a Nikon D60 camera (experiments C-E). All leaf sections were assessed for both hyphal coverage and number of incidences of penetration.

In experiments A-C, the hyphal coverage and incidence of penetration was assessed on a binary scale (i.e. a score of "0" was assigned if the trait was absent, and a score of "1" if it was present). In experiments D and E, the total area of each leaf segment was calculated by measuring its length and width at 100x magnification using an eyepiece micrometer. The hyphal coverage was expressed as a percent of leaf tissue obscured by hyphae. For the incidence of penetration, the total number of penetration events on each leaf was counted and divided by the total area of the leaf segment for intra-experimental comparisons.

6.2.5 Statistical methods

Statistical analyses were conducted using the results from experiments D and E. For all statistical tests, a type 1 error threshold of 0.05 was used. All statistical analyses were completed
using SAS v. 9.1 (SAS Institute Inc., Cary, NC, USA). Data from each experiment were analysed separately. A non-inoculated "blank," treated with water, was included at each time point in both experiments. There were three replicates for each treatment condition. The fixed effects were the strain of inoculum applied, the identity of the host plant, and the date of collection.

The data were sorted by date of collection and the host plant of origin, then assessed for normality using the Shapiro-Wilk test \( (H_0: \text{the samples come from a normally distributed population}) \), and were tested for homogeneous variance using the Levene test \( (H_0: \text{the variances of the data are equal}) \). For data sets that were not normally distributed or that did not exhibit homogenous variance, the non-parametric Wilcoxon rank-sum test (also known as the Mann-Whitney-U test; \( H_0: \text{the samples come from identical populations} \) ) was used to compare the means within each time point.

A general linear model by date of collection and identity of the host plant was utilized. The classification variables were the fungal inoculum's species, host-plant origin, and isolate number. The model was that the number of penetration incidences per unit area were dependent on the isolate number. The means and least significance difference (LSD) were calculated for each isolate. An example of the SAS statements used in these analyses may be found in Appendix 6.1.

### 6.3 Results

#### 6.3.1 Experiment A

The goal of experiment A was to observe the infection of wheat and Kentucky bluegrass leaves by *M. nivale* and *M. majus*. In this experiment, hyphal plugs were the sole source of inoculum used, and two isolates of each species were used (Figure 6.2). Under these conditions, penetration of the leaf tissue was first observed at 2 days post-inoculation (dpi) on both host
plants. On Kentucky bluegrass, both isolates of both *M. nivale* and *M. majus* penetrated the leaf tissue at 2 dpi, whereas on wheat, only the *M. nivale* isolates (both originally collected from grass) penetrated the leaf tissue at 2 dpi. Penetration of wheat leaves by *M. majus* was first observed at 3 dpi.

The mechanism of penetration was the same on both leaf types and for both *M. nivale* and *M. majus*: hyphae were observed growing directly into the stomata (Figure 6.3). When the leaves were heavily infected, hyphae were also observed growing out of the stomata (Figure 6.4). Neither appressoria nor haustoria were observed on any leaf type or with either of the fungal species tested.

### 6.3.2 Experiment B

The goal of experiment B was to determine whether the timing and/or the mechanism of penetration may differ when a conidial suspension, rather than a hyphal suspension, was used to apply inoculum to the leaf sections. Two isolates of *M. majus* and four isolates of *M. nivale* were included in this experiment, where two of the *M. nivale* isolates were originally collected from wheat, and the other two were collected from turfgrass. In addition, a suspension containing hyphal tissue, rather than a plug of agar (as in experiment A) was used to apply the hyphal inoculum.

For leaf segments inoculated with a hyphal suspension (Figure 6.5), similar trends were observed relative to those obtained in experiment A. On Kentucky bluegrass, five of the isolates, with the exception of one of the *M. nivale* isolates originally from turfgrass, penetrated the leaf tissue at 2 dpi. On the wheat leaves, only one of the *M. majus* and one of the *M. nivale* isolates from wheat had penetrated the leaf tissue by 2 dpi. At 3 dpi, both of the *M. majus* isolates, one of
the two *M. nivale* wheat isolates, and both of the turf isolates were observed penetrating the Kentucky bluegrass leaf segments, and all six of the isolates had penetrated the wheat leaves. Between experiments A and B, the application of hyphal inoculum by using a hyphal suspension rather than an agar plug yielded nearly identical results.

For leaf segments inoculated with a conidial suspension (Figure 6.6), penetration was delayed by one day relative to the leaves inoculated with hyphae. On Kentucky bluegrass, only the *M. majus* isolates were observed penetrating the leaf tissue at 3 dpi, whereas *M. nivale* was not observed penetrating the leaf tissue until 4 dpi. On wheat, the *M. nivale* isolate was never observed penetrating the leaf tissue. Only a single *M. majus* isolate had penetrated the leaf tissue at 3 dpi, and the incidence of penetration remained infrequent and inconsistent for both *M. majus* isolates until 6 dpi. As in experiment A, the sole mechanism of penetration observed was the direct penetration of stomata. Neither appressoria nor haustoria were observed.

### 6.3.3 Experiment C

Experiment C was similar to experiment B, but in this experiment, the plates were incubated in the dark, rather than under full light. Both hyphal and conidial suspensions were used to inoculate detached leaf segments of wheat and Kentucky bluegrass. On Kentucky bluegrass inoculated with a hyphal suspension (Figure 6.7), both of the *M. majus* isolates were observed penetrating the leaf tissue at 1 dpi, whereas penetration for *M. nivale* was not observed until 2 dpi. On wheat inoculated with a hyphal suspension (Figure 6.7), penetration was observed for only a single *M. majus* isolate at 1 dpi. The second *M. majus* isolate was observed penetrating the leaf tissue at 2 dpi, and the *M. nivale* isolate penetrated the tissue at 3 dpi.
When wheat and Kentucky bluegrass were inoculated with conidia from two isolates of *M. nivale*, both originally collected from turfgrass (Figure 6.8), penetration was not observed until 3 dpi for either leaf type. As in experiments A and B the sole mechanism of penetration observed in experiment C was the direct penetration of stomata. Neither appressoria nor haustoria were observed.

**6.3.4 Experiment D**

In experiment D, the only inoculum type applied was a hyphal suspension. The inoculated leaf sections were incubated in the dark at 4 °C, rather than at room temperature as in previous experiments. In addition, the method by which the presence or absence of hyphal colonization and penetration was recorded was changed relative to the previous experiments (Section 6.24).

In this experiment, the number of incidences of penetration by unit area on the two host plants and for each day were compared between the isolates, according to their species and their host plant of origin. Two isolates of *M. majus* and four of *M. nivale* (two from wheat and two from turf) were included in this experiment.

On Kentucky bluegrass (Figure 6.9), penetration was first observed by one of the *M. nivale* isolates from turf at 3 dpi. All six of the isolates included in this study penetrated the leaf tissue by 6 dpi. On wheat (Figure 6.9), one *M. majus* isolate penetrated the leaf tissue at 4 dpi. All of the isolates, with the exception of one *M. majus* isolate and one *M. nivale* isolate from turf, had penetrated the leaf tissue by 6 dpi.

Most of the data for each isolate's incidence of penetration at each date and between the two host plants were found to be non-normally distributed (p-value for Shapiro-Wilk test > 0.05) and/or exhibited non-homogeneous variance (p-value for Levine test < 0.05). For this reason, the
statistical significance of these data was assessed using the non-parametric Wilcoxon two-sample test, using the two-sided t approximation (Pr > |Z|) to identify significant differences in the mean number of penetration events between the two leaf types for each individual isolate (Table 6.4). Using this test, none of the data were found to be statistically significant.

When the data from this experiment were pooled by the host origin of the fungal isolates, a different pattern emerged (Table 6.5). Because the data for the number of incidences of penetration on the two host plants were found to be non-normally distributed (p-value for Shapiro-Wilk test > 0.05) and/or exhibited non-homogeneous variance (p-value for Levine test < 0.05), once again the Wilcoxon test was employed to identify significant differences within a host origin group. When the data were partitioned in this manner, the mean incidences of penetration differed significantly between the turf and wheat isolates on both Kentucky bluegrass and on wheat at 6 and 7 dpi: in both cases, a larger number of penetration occurrences were observed for the turf-derived isolates than the wheat-derived isolates on both host plants.

6.3.5 Experiment E

The purpose of experiment E was to repeat the previous room-temperature experiments using a hyphal suspension but with the updated scoring system (experiment D) to record the incidences of leaf colonization and penetration (Figure 6.10). On Kentucky bluegrass, penetration of the leaf tissue was observed at 2 dpi for leaf segments inoculated with one of the M. majus isolates, one M. nivale turf isolate, and one M. nivale wheat isolate. At 3 dpi, all isolates with the exception of one M. nivale wheat isolate were observed penetrating the leaf tissue. On wheat, one M. majus isolate and one M. nivale isolate from wheat were observed penetrating the leaf tissue at 2 dpi.
At 3 dpi, all of the isolates with the exception of one of the *M. nivale* wheat isolates had penetrated the leaf tissue.

In experiment E, non-normality and/or non-homogenous variance were found for both the by-isolate and the by-host-origin datasets, so the data were analysed as described for experiment D. In experiment E, none of the means in either analysis were found to differ significantly (Table 6.6 and Table 6.7).

### 6.4 Discussion

Several independent experiments were conducted to elucidate the physical mechanisms and the timing of the infection processes of *M. nivale* and *M. majus* on the representative host plants *T. aestivum* and *P. pratensis*. The influences of inoculum type (conidial vs. hyphal), incubation temperature (4 °C vs. 23 °C), and host plant of origin (wheat vs. turfgrass) were investigated by observing the rate of fungal colonization of the plant tissue and by counting the incidences of penetration into the plant tissue. All of the isolates used in these experiments displayed approximately the same rate of growth on PDA at room temperature, which is correlated to virulence (Hofgaard et al. 2006).

Penetration of hyphae into the plant tissue was observed in every experiment conducted. Neither appressoria nor haustoria were observed in any experiment or on either host plant; instead, the hyphae were observed growing directly into the stomata without the formation of specialized structures (Figure 6.3). In addition, when the plant tissue was heavily colonized, hyphae were observed growing out of the stomata (Figure 6.4). These observations are in line with early descriptions of the infection process (Dahl 1934). However, two recent studies, both using the same single strain of *M. nivale* (originally collected from rye) to prepare all of the
inoculum, examined the infection process of *M. nivale* on the hybrid cereal *Triticale* and on rye (*Secale cereale*) and suggested that this pathogen forms haustoria or haustorium-like structures within the plant tissue (Dubas et al. 2011; Zur et al. 2011). The lack of haustoria in the experiments reported herein may be due to the use of detached senescing leaves in the experiments described in this thesis rather than attached vigorous leaves as were used in the experiments by Dubas et al. (2011) and by Zur and colleagues (2011). Biotrophy may be an important strategy for these pathogens when the plant is actively growing, but when the leaves are inactive or dead, as in these experiments, saprotrophy may be employed instead.

One goal of the experiments described in this Chapter was to explore the claim that conidia alone do not cause infection on plant tissue. In both experiments that included conidial inoculum, the leaves inoculated with conidial suspensions displayed a much lower rate of colonization relative to leaves inoculated with hyphae and incubated under the same conditions. These same conidial suspensions were known to contain living spores capable of normal growth because aliquots of the spore suspensions used in all experiments were incubated on artificial media and were observed to germinate and to produce colonies identical to those established using actively-growing hyphae. Despite this broad similarity, the experiments including conidia revealed large differences between the different strains used. In experiment B, the conidia of the single *M. nivale* isolate studied were not observed to germinate and to subsequently infiltrate the wheat tissue within the timeframe of the experiment; however, in experiment C, both of the *M. nivale* isolates studied penetrated the leaf tissue of wheat by 4 dpi. While one of the two *M. nivale* isolates used in experiment C was observed penetrating the wheat tissue at 3 dpi, and the observation of penetration remained consistent for the remaining duration of the experiment, the other isolate studied was first observed penetrating the wheat at 4 dpi, and was not observed
penetrating the leaf tissue consistently until 6 dpi. The *M. majus* isolates included in experiment B also yielded inconsistent results, and also failed to consistently penetrate the leaf tissue until 6 dpi.

On Kentucky bluegrass, the *M. nivale* isolate included in experiment B yielded inconsistent results. Penetration of the leaf tissue was first observed at 4 dpi, but was not observed again until 7 dpi. In contrast, both of the *M. nivale* isolates included in experiment C consistently penetrated the leaf tissue from 3 dpi through to the end of the observation period. The *M. majus* conidia used in experiment B consistently penetrated the tissue of Kentucky bluegrass every day after 3 dpi.

The inconsistent observation of penetration for most of the isolates included in these studies, in addition to the failure of conidia to produce rapid hyphal coverage on the surface of the plant leaves, suggests that the surface of the plant is a less favourable environment for spore germination than artificial media. Because detached leaf sections, rather than whole and / or living leaves were used in these experiments, it is unlikely that this observation can be explained by the activation of the plant's active defences. Instead, a more likely explanation is that the properties of the leaves' surfaces, such as the waxy cuticle, may have prevented the conidia from adhering long enough to germinate. This suggestion is corroborated by the fact that "stray" ungerminated conidia were not observed on the leaf surface at any time.

Together, the observations that conidia are capable of rapid germination and growth on artificial media, but apparently not on plant tissue, corroborates earlier claims that the conidia of *M. nivale* and *M. majus* do not cause plant disease directly. Instead, conidia may be important in the dispersal of these pathogens. After dispersal, the conidia could germinate in the thatch or on
dead plant tissue and grow saprotrophically, producing a mycelial mass that may later cause infection of the plant when conditions become favourable for the fungus.

In all experiments, the hyphal inoculum of both M. nivale and M. majus, regardless of the host species of origin, colonized and penetrated the leaf tissue of both of the host species studied, P. pratensis and T. aestivum. Regardless of whether it was delivered via a plug of agar or by a hyphal suspension, hyphal inoculum resulted in the colonization and penetration of the leaf tissue within 48 hours post-inoculation in most cases. Penetration was observed by 72 hours post-inoculation in most cases when conidial inoculum was used. Together, these experiments suggest that both M. nivale and M. majus, regardless of the host plant from which they were originally isolated, are capable of penetrating the leaf tissue of both turfgrasses and cereals. The observed host preferences in the field may thus reflect environmental and biotic factors that might limit spread.

Despite the overall trends in the timing of infection as described above, some inconsistencies were observed within and between experiments. Although a hyphal suspension was used in both experiments C and E, and the leaf tissue was incubated at the same temperature in both experiments, the timing of penetration by M. majus isolate 99049 was delayed by 1 day on both leaf types in experiment C relative to experiment E. In contrast, for the M. nivale turf isolates, the penetration of Kentucky bluegrass was first observed at 2 dpi in both experiments E and C, but at 3 dpi in experiment C and at 2 dpi in experiment E. These conflicting observations suggest that, although there were no consistent differences between the observations for the isolates included in the experiments described, the trends observed for individual isolates were variable. These differences underscore the necessity for these observations to be interpreted as general
trends in the timing when penetration may be expected, and also the need for a greater number of replications so that stochastic variation can be taken into account.

Among the *M. nivale* isolates studied, the host species from which the isolate was originally collected did not exhibit a strong influence on the isolate's ability to penetrate the tissues of either of the host plants under study. This was unexpected, because genotypic differences have been identified between populations of *M. nivale* growing on different turfgrass species (Mahuku et al. 1998), and *M. nivale* isolates from wheat are genetically distinct from those originating from turf (see Chapters 2 and 3). Although *M. majus* has only been isolated from cereals, and has not been reported on turfgrasses (Maurin et al. 1995), this pathogen is clearly capable of colonizing and penetrating the tissue of at least one turfgrass species under artificial inoculation conditions. One explanation for the apparent lack of *M. majus* on grasses in the field may be that *M. nivale* or other phylloplane organisms may simply out-compete *M. majus* on turfgrasses, perhaps due to differences in their abilities to either trigger or resist plant defensive processes. For example, *M. majus* has been shown to out-compete *M. nivale* on rye (*Secale cereale*) (Hofgaard et al. 2006), and also displays a decreased sensitivity to the plant defensive compound benzoxazolinone, which is produced in large quantities by rye (Simpson et al. 2000). A more detailed analysis of the plant's responses to the presence of these pathogens (e.g. the identification and monitoring of transcript levels of defensive enzymes) may provide further insight into the relationships between these pathogens and their preferred hosts.

In addition to the experiments conducted at room temperature, the inoculated leaves in experiment D were incubated at 5 °C. A lower temperature was chosen to simulate the conditions in the field at which *M. nivale* and *M. majus* typically cause disease. At the cooler temperature, the timing of surface colonization and penetration was delayed, but not prevented. The trends
observed, with respect to the slightly earlier penetration by pathogens on their native hosts, were unchanged. This delay in the penetration of plant tissues is in line with the slower growth rate of these pathogens on artificial media at these lower temperatures (Snider et al. 2000). Despite this reduced growth rate, the ability of these pathogens to remain active at low temperatures (i.e. under snow or between 0-15 °C) may thus allow them to exploit their host plants when other pathogens are inactive. For this reason, it may be possible that *M. nivale* and *M. majus* cause more frequent disease at cooler rather than warmer temperatures simply because they are out-competed by other pathogens in the field.

Taken together, the results presented in this chapter suggest that there are no obvious differences in the mechanisms by which *M. nivale* and *M. majus* attack their host plants. Both species were capable of colonizing and penetrating plant tissue from both their native- and non-native hosts, and the temperature of incubation delayed but did not otherwise alter these patterns. Hyphal inoculum resulted in the rapid colonization of plant tissue, whereas conidial inoculum yielded inconsistent results, but generally resulted in a slower rate of tissue colonization.
6.5 References for Chapter 6

Pronczuk, M., and Messysz, M. 1991. Infection ability of mycelium and spores of Microdochium nivale (Fr.) Samuels & Hallett to Lolium perenne L. Mycotoxin Research 7A.


Table 6.1 Summary of conditions tested in infection process experiments performed for *M. nivale* and *M. majus* inoculated on *P. pratensis* and *T. aestivum*.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Inoculum Types*</th>
<th>Number of Isolates</th>
<th>Incubation Temperature (°C)</th>
<th>Light condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>HP</td>
<td>2</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>B</td>
<td>HP, CS</td>
<td>3</td>
<td>5</td>
<td>23</td>
</tr>
<tr>
<td>C</td>
<td>HS, CS</td>
<td>1</td>
<td>3</td>
<td>23</td>
</tr>
<tr>
<td>D</td>
<td>HS</td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>E</td>
<td>HS</td>
<td>3</td>
<td>3</td>
<td>23</td>
</tr>
</tbody>
</table>

* HP = hyphal plug; CS = conidia suspension; HS = hyphal suspension

† 50 µmol/m²/s for 24 h each day
<table>
<thead>
<tr>
<th>Isolate</th>
<th>Species</th>
<th>Host Origin</th>
<th>Geographic Origin</th>
<th>Experiment and Inoculum Type*</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>99049</td>
<td><em>M. majus</em></td>
<td><em>T. aestivum</em></td>
<td>Atwood, ON</td>
<td></td>
<td>HP</td>
<td>HS, CS</td>
<td>HS</td>
<td>HS</td>
<td>HS</td>
</tr>
<tr>
<td>99061</td>
<td><em>M. majus</em></td>
<td><em>T. aestivum</em></td>
<td>Atwood, ON</td>
<td></td>
<td>HP</td>
<td>HS</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10148</td>
<td><em>M. majus</em></td>
<td><em>T. aestivum</em></td>
<td>France</td>
<td></td>
<td>-</td>
<td>CS</td>
<td>HS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12045</td>
<td><em>M. majus</em></td>
<td><em>T. aestivum</em></td>
<td>Ottawa, ON</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>HS</td>
</tr>
<tr>
<td>12430</td>
<td><em>M. majus</em></td>
<td><em>T. aestivum</em></td>
<td>Ottawa, ON</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>HS</td>
<td>HS</td>
</tr>
<tr>
<td>10085</td>
<td><em>M. nivale</em></td>
<td><em>P. annua</em></td>
<td>Guelph, ON</td>
<td></td>
<td>HP</td>
<td>CS</td>
<td>HS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10086</td>
<td><em>M. nivale</em></td>
<td><em>P. annua</em></td>
<td>Guelph, ON</td>
<td></td>
<td>HP</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>11011</td>
<td><em>M. nivale</em></td>
<td><em>P. pratensis</em></td>
<td>Guelph, ON</td>
<td></td>
<td>-</td>
<td>HS</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>11016</td>
<td><em>M. nivale</em></td>
<td><em>P. pratensis</em></td>
<td>Guelph, ON</td>
<td></td>
<td>-</td>
<td>-</td>
<td>CS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>11228</td>
<td><em>M. nivale</em></td>
<td><em>P. annua</em></td>
<td>Guelph, ON</td>
<td></td>
<td>-</td>
<td>HS</td>
<td>CS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>99084</td>
<td><em>M. nivale</em></td>
<td><em>T. aestivum</em></td>
<td>Atwood, ON</td>
<td></td>
<td>-</td>
<td>HS</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>99052</td>
<td><em>M. nivale</em></td>
<td><em>T. aestivum</em></td>
<td>Atwood, ON</td>
<td></td>
<td>-</td>
<td>HS</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12260</td>
<td><em>M. nivale</em></td>
<td><em>T. aestivum</em></td>
<td>Ottawa, ON</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>HS</td>
<td>HS</td>
</tr>
<tr>
<td>12262</td>
<td><em>M. nivale</em></td>
<td><em>T. aestivum</em></td>
<td>Ottawa, ON</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>HS</td>
<td>HS</td>
</tr>
<tr>
<td>12267</td>
<td><em>M. nivale</em></td>
<td><em>P. pratensis</em></td>
<td>Guelph, ON</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>HS</td>
<td>HS</td>
</tr>
<tr>
<td>12438</td>
<td><em>M. nivale</em></td>
<td><em>P. pratensis</em></td>
<td>Guelph, ON</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>HS</td>
<td>-</td>
</tr>
</tbody>
</table>

*HP = hyphal plug; HS = hyphal suspension; CS = conidia suspension; - = not included*
Table 6.3 Sample collection timepoints for infection process studies of *M. nivale* and *M. majus* on *T. aestivum* and *P. pratensis*.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Collection times*</th>
<th>Hours post-inoculation</th>
<th>Days post-inoculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td>6, 8, 10, 12, 24</td>
<td>2, 3, 4, 5, 6, 7</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>-</td>
<td>1, 2, 3, 4, 5, 6, 7</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td>12, 14, 16, 18, 20, 24</td>
<td>2, 3, 4, 5, 6, 7</td>
</tr>
<tr>
<td>D</td>
<td></td>
<td>-</td>
<td>1, 2, 3, 4, 5, 6, 7</td>
</tr>
<tr>
<td>E</td>
<td></td>
<td>-</td>
<td>1, 2, 3, 4, 5, 6, 7</td>
</tr>
</tbody>
</table>

* All collection times that occurred within the first 24 hours following inoculation are listed in the hours post-inoculation column; all other collection times are listed in the days post-inoculation column.
Table 6.4 Comparisons between mean number of penetration observations per unit area for each isolate of Microdochium nivale (Mn) and M. majus (Mm) on detached leaves of Kentucky bluegrass (K) and wheat (W) in experiment D. Data for each isolate at each time point were compared using a two-sided Wilcoxon rank sum approximation. Means followed by different letters were significantly different at p = 0.05.

<table>
<thead>
<tr>
<th>Days post-inoculation</th>
<th>12260 (MnW)</th>
<th>12262 (MnW)</th>
<th>12267 (MnK)</th>
<th>12438 (MnK)</th>
<th>12430 (MmW)</th>
<th>99049 (MmW)</th>
<th>H2O</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K</td>
<td>W</td>
<td>K</td>
<td>W</td>
<td>K</td>
<td>W</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
</tr>
<tr>
<td>2</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
</tr>
<tr>
<td>3</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
<td>41a</td>
<td>0a</td>
<td>0a</td>
</tr>
<tr>
<td>4</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
<td>22a</td>
<td>0a</td>
</tr>
<tr>
<td>5</td>
<td>0a</td>
<td>0a</td>
<td>24a</td>
<td>4a</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
</tr>
<tr>
<td>6</td>
<td>357a</td>
<td>11a</td>
<td>11a</td>
<td>0a</td>
<td>1009a</td>
<td>55a</td>
<td>231a</td>
</tr>
<tr>
<td>7</td>
<td>60a</td>
<td>13a</td>
<td>532a</td>
<td>56a</td>
<td>3973a</td>
<td>180a</td>
<td>3887a</td>
</tr>
</tbody>
</table>


**Table 6.5** Comparisons between mean number of penetration observations per unit area for all isolates, regardless of identity, from each host type on Kentucky bluegrass and wheat in experiment D. Data for each host type at each time point were compared using a two-sided Wilcoxon rank sum approximation. Means followed by different letters were significantly different at $p = 0.05$.

<table>
<thead>
<tr>
<th>Days post-inoculation</th>
<th>Kentucky bluegrass</th>
<th>Wheat</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>turf-derived isolates</td>
<td>wheat-derived isolates</td>
</tr>
<tr>
<td>1</td>
<td>0a</td>
<td>0a</td>
</tr>
<tr>
<td>2</td>
<td>0a</td>
<td>0a</td>
</tr>
<tr>
<td>3</td>
<td>21a</td>
<td>0b</td>
</tr>
<tr>
<td>4</td>
<td>14a</td>
<td>0a</td>
</tr>
<tr>
<td>5</td>
<td>345</td>
<td>6a</td>
</tr>
<tr>
<td>6</td>
<td>620a</td>
<td>216b</td>
</tr>
<tr>
<td>7</td>
<td>3930a</td>
<td>217b</td>
</tr>
</tbody>
</table>
Table 6.6 Comparisons between mean number of penetration observations per unit area for each isolate of *Microdochium nivale* (Mn) or *M. majus* (Mm) on detached leaves of Kentucky bluegrass (K) and wheat (W) in experiment E. Data for each isolate at each time point were compared using a two-sided Wilcoxon rank sum approximation. Means followed by different letters were significantly different at $p = 0.05$.

<table>
<thead>
<tr>
<th>Days post-inoculation</th>
<th>12260 (MnW)</th>
<th>12262 (MnW)</th>
<th>12267 (MnK)</th>
<th>12430 (MnK)</th>
<th>12045 (MmW)</th>
<th>99049 (MnW)</th>
<th>H₂O</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K</td>
<td>W</td>
<td>K</td>
<td>W</td>
<td>K</td>
<td>W</td>
<td>K</td>
</tr>
<tr>
<td>1</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
<td>0a</td>
</tr>
<tr>
<td>2</td>
<td>0a</td>
<td>0a</td>
<td>7a</td>
<td>3a</td>
<td>1a</td>
<td>0a</td>
<td>0a</td>
</tr>
<tr>
<td>3</td>
<td>0a</td>
<td>0a</td>
<td>230a</td>
<td>49a</td>
<td>42a</td>
<td>4a</td>
<td>190a</td>
</tr>
<tr>
<td>4</td>
<td>1a</td>
<td>0a</td>
<td>150a</td>
<td>45a</td>
<td>69a</td>
<td>9a</td>
<td>349a</td>
</tr>
<tr>
<td>5</td>
<td>270a</td>
<td>0a</td>
<td>273a</td>
<td>63a</td>
<td>166a</td>
<td>63a</td>
<td>367a</td>
</tr>
<tr>
<td>6</td>
<td>357a</td>
<td>3a</td>
<td>653a</td>
<td>156a</td>
<td>379a</td>
<td>60a</td>
<td>297a</td>
</tr>
<tr>
<td>7</td>
<td>646a</td>
<td>101a</td>
<td>275a</td>
<td>58a</td>
<td>572a</td>
<td>242a</td>
<td>466a</td>
</tr>
</tbody>
</table>

414
Table 6.7 Comparisons between mean number of incidences of penetration per unit area for all isolates from each host type on detached leaves of Kentucky bluegrass and wheat in experiment E. Data for each host type at each time point were compared using a two-sided Wilcoxon rank sum approximation. Means followed by different letters were significantly different at $p = 0.05$.

<table>
<thead>
<tr>
<th>Days post-inoculation</th>
<th>Kentucky bluegrass host</th>
<th>Wheat host</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>turf-derived isolates</td>
<td>wheat-derived isolates</td>
</tr>
<tr>
<td></td>
<td>wheat-derived isolates</td>
<td>turf-derived isolates</td>
</tr>
<tr>
<td>1</td>
<td>0a</td>
<td>0a</td>
</tr>
<tr>
<td>2</td>
<td>1a</td>
<td>1a</td>
</tr>
<tr>
<td>3</td>
<td>25a</td>
<td>2a</td>
</tr>
<tr>
<td>4</td>
<td>33a</td>
<td>6a</td>
</tr>
<tr>
<td>5</td>
<td>67a</td>
<td>44a</td>
</tr>
<tr>
<td>6</td>
<td>113a</td>
<td>103a</td>
</tr>
<tr>
<td>7</td>
<td>158a</td>
<td>156a</td>
</tr>
</tbody>
</table>
Figure 6.1 The disease cycle describing the events that occur during a host-pathogen interaction (modified from (Agrios 2005))
Figure 6.2 Number of detached leaf segments of *P. pratensis* (A) and *T. aestivum* (B) where penetration was observed at time of collection for leaf blades treated with hyphal inoculum in experiment A. Three leaf blades were collected at each time point.
Figure 6.3 Penetration of stomata of detached leaves of *T. aestevum* by hyphae of *M. majus* isolate 99049 (stained blue) incubated on moistened filter paper and incubated at 22 ºC. Photo taken at 3 dpi, 400 x magnification.
Figure 6.4 Hyphae of *M. majus* isolate 99061 emerging from stomata of detached leaves of wheat (circled) incubated on moistened filter paper and incubated at 22 °C. Photo taken at 400x magnification, 5 dpi.
Figure 6.5 Number of detached leaf segments of *P. pratensis* (A) and *T. aestivum* (B) incubated on moist filter paper where penetration was observed at time of collection for leaf blades treated with hyphal inoculum in experiment B. Three leaf blades were collected at each time point.
Figure 6.6 Number of detached leaf segments of *P. pratensis* (A) and *T. aestivum* (B) incubated on moist filter paper where penetration was observed at time of collection for leaf blades treated with conidial inoculum in experiment B. Three leaf blades were collected at each time point.
Figure 6.7 Number of detached leaf segments of *P. pratensis* (A) and *T. aestivum* (B) incubated on moist filter paper where penetration was observed at time of collection for leaf blades treated with hyphal inoculum in experiment C.
Figure 6.8 Number of detached leaf segments of *P. pratensis* (A) and *T. aestivum* (B) incubated on moist filter paper where penetration was observed at time of collection for leaf blades treated with conidial inoculum in experiment C.
Figure 6.9 Number of incidences of penetration per unit area on *P. pratensis* (A) and *T. aestivum* (B) incubated on moist filter paper at time of collection for leaf blades treated with hyphal inoculum in experiment D.
Figure 6.10 Number of incidences of penetration per unit area on *P. pratensis* (A) and *T. aestivum* (B) incubated on moist filter paper at time of collection for leaf blades treated with hyphal inoculum in experiment E.
Appendices for Chapter 6

Appendix 6.1: SAS Statements

* FILENAME: 130521_expE.sas;
* DATE: 13.05.21;

* Data from infection study experiment;
* Precede comments with an asterisk, and end with semicolon;

data temp;
options pagesize=200 linesize=120;
infile cards;
* dlm='09'x;
* dlm is delimiter and '09'x is the ascii symbol for tabs;
input day $ isolate $ species $ origin $ host $ leaf_no pen_by_area;
cards;
1 12045 M W K 1 0
1 12045 M W K 2 0
...
;
run;

proc sort; by day host;
proc univariate normal; by day host; var pen_by_area;
run;
proc npar1way; by day host;
class origin;
run;

proc glm; by day host;
class species isolate origin;
model pen_by_area = origin;
means origin/ LSD;
means origin / hovtest;
run;

proc sort; by day isolate;
proc glm; by day isolate;
class species origin host;
model pen_by_area = host;
means host/ LSD;
means host / hovtest;
run;
proc univariate normal; by day isolate; var pen_by_area;
run;
proc npar1way; by day isolate;
class host;
run;
Chapter 7 General Discussion and Conclusions

7.1 Major conclusions

The major conclusions from the research presented in this thesis are as follows:

1. *Microdochium nivale* and *M. majus* exhibit consistent genetic and genomic differences that validate their status as distinct species (Chapters 2 and 3). The genera *Microdochium* and *Monographella* (in which these species have been placed) have not yet been reviewed under the one fungus-one name system (Taylor 2011). Because the names based on the asexual stage are older (have priority), are better known, and are of less uncertain application than the names based on the sexual stage, they have been used in this thesis.

2. Intraspecific genetic differences exist within *M. nivale*, both within isolates collected from a single type of host plant and between isolates collected from different species (Chapters 3 and 4). These findings are in agreement with previous reports that *M. nivale* has a high level of genetic variation.

3. The formation of ascospores (and hence the implication of sexual reproduction) was induced in the lab for *M. majus*, but could not be confirmed in *M. nivale*. The mating type genes in both species and in other members of the same order failed to conform consistently to the syntenic arrangement of this genetic region that was predicted based on other fungi. In addition, only the MAT1-2-1 mating gene was detected among all of the *Microdochium* and *Xylariales* strains examined. As a result, although *M. majus* displayed apparently homothallic mating, it was not possible to label *M. nivale* as homothallic, heterothallic, or pseudohomothallic (Chapter 5).
4. Despite host preferences reported from field studies, infection process studies on detached leaves did not reveal pathogenic differences between *M. nivale* and *M. majus* and between isolates collected from different plant hosts. When infection occurred, neither appressoria nor haustoria were observed; however, this observation may have been influenced by the use of detached leaves, rather than whole plants (Chapter 6).

7.2 General Discussion and Conclusions

The major goal of this thesis was to investigate differences between *Microdochium nivale* and *M. majus*, two fungal plant pathogens that were recently recognized as distinct species (Glynn et al. 2005) rather than conspecific varieties. This nomenclatural change was preceded by claims that the recognition of *M. nivale* and *M. majus* as separate varieties was not meaningful based on morphological observations (Litschko and Burpee 1987). As a result, the projects described herein were undertaken to explore this apparent contradiction. At the beginning of the research project (Chapter 2), genetic differences between *M. nivale* and *M. majus* were investigated by examining the nucleotide sequences of four genomic regions for a small sample of isolates from various host plant and geographic origins. Among the four regions examined, only ITS failed to resolve the species into distinct clades; the remaining genetic regions examined confirmed that *M. nivale* and *M. majus* are genetically distinct. Further divisions were resolved within *M. majus*, where the isolates from Europe and North America were grouped separately by the analysis of RPB2, and within *M. nivale*, for which the sequences of β-tubulin and RPB2 suggested that isolates obtained from wheat are genetically distinct from turfgrass isolates.
The findings of Chapter 2 were further investigated by using next-generation sequencing technology to produce draft genome sequences for several isolates of *M. nivale* and *M. majus* and one isolate of *M. bolleyi*, a root-inhabiting species (Chapter 3). When these whole-genome sequences were compared within and between species, several interesting trends were observed. Despite originating from North America and Europe, respectively, the genomes of two *M. majus* isolates were found to be nearly identical in terms of the number and the identities of their predicted gene sequences. A similar trend was observed for the two *M. nivale* wheat-derived isolates sequenced, which were also collected from two different continents. In contrast, the turf isolate of *M. nivale* exhibited genome-wide differences relative to the wheat isolates. The *M. nivale* turf isolate shared between 78-79% of its predicted genes with the other *M. nivale* and *M. majus* isolates at an e-value of 1e-50, the two *M. nivale* wheat isolates shared about 92% of their predicted genes, and the two *M. majus* isolates shared 95% of their predicted genes. When a tree was constructed using 10 randomly selected sequences from *Microdochium*, the turf isolate was found to be as dissimilar from the other *M. nivale* and *M. majus* sequences as it was from *M. bolleyi*. The biotypes of *M. nivale* and *M. majus* on different host plants appear to be genetically distinct, and gene exchange may be limited between isolates from different host species. This finding is broadly congruent with reports that the population of *M. nivale* on turfgrass may be distinct from that of cereals (Smith 1983).

*Microdochium nivale* and *M. majus* exhibit different sensitivities to plant defensive enzymes (Simpson et al. 2000), which is likely related to their host preferences and may prevent their interaction in the field. The presence of different sensitivities to host defensive enzymes or compounds within either species is unknown, but is a potential area for future investigation. In addition, the observation that the genomes of *M. nivale* isolates displayed a lower level of
similarity with one another than did the genomes of *M. majus* isolates is also in agreement with studies that examined RAPD patterns (Lees et al. 1995) as well as RFLP patterns of ITS and esterase protein profiles (Maurin et al. 1995), which revealed that *M. nivale* isolates displayed a higher level of variability than *M. majus*.

However, the suggestion that individual isolates may exhibit strong host specialization contrasts with the observations in Chapter 6 that all isolates of *M. nivale* and *M. majus* examined were capable of causing infection on both *Triticum aestivum* and *Poa pratensis* regardless of their host origin. This observation is in agreement with observations of *M. nivale* and *M. majus* on wheat seedlings, wherein host origin and fungal species were not strongly correlated with aggressiveness (Maurin et al. 1995). The observations of the physical infection process in this thesis may have been affected by the conditions used in these experiments: by inoculating spore suspensions of single fungal isolates on detached plant leaves under sterile conditions, and the lack of competition by other microbes may have permitted fungi to invade successfully where they would have failed in the field. Similarly, when the infection process experiments were performed at 23 °C compared to 4 °C, the timing of the infection was delayed, but otherwise the physical process of infection was unchanged in either case. This suggests that the ability of these pathogens to cause disease at low temperatures in the field does not preclude them from infecting plant tissue at warmer temperatures; instead, a more likely explanation for their scarcity as disease-causing agents at higher temperatures is likely the presence of microbial competitors or the increased activity of plant defences. In co-inoculation studies, host preferences have been observed between *M. nivale* and *M. majus* on oat (*Avena sativa*), barley (*Hordeum vulgare*), and wheat *T. aestivum* (Simpson et al. 2000). However, when *M. nivale*, *M. majus*, and *Fusarium culmorum* (a wheat pathogen) were inoculated on wheat or rye, the colonization success of each
species was dependent on both the presence or absence of other pathogen species as well as the temperature of inoculation (Simpson et al. 2004).

Throughout the infection process experiments, neither appressoria nor haustoria were observed. This contrasts with reports that haustoria were produced during the infection of *M. nivale* on the hybrid cereal triticale (*x Triticosecale*) and on rye (*Secale cereale*) (Dubas et al. 2011; Zur et al. 2011). Aside from possible host species differences, this may also be related to the use of detached leaves in the current experiments rather than whole plants (as done in the cited studies). This finding is important when contrasted with an earlier report that, at least in terms of isolate aggressiveness, detached leaf assays provided a meaningful proxy for experiments with whole plants (Diamond and Cooke 1999). To investigate the importance of competition in the host- and temperature-specificities observed in the field, experiments using whole plants may help to clarify this apparent contradiction. It may also be valuable to prepare fluorescently labeled strains of *M. nivale* and *M. majus*, which would facilitate the monitoring of the infection process of these pathogens in the presence of microbial competitors. Although attempts were made to create such strains, they were ultimately unsuccessful and were not reported in this thesis.

Another tool to investigate genetic diversity in relation to host-specific populations is the examination of microsatellite regions by ISSR and / or SSR. Within this thesis, this resource was applied to isolates of *M. nivale* collected from two proximate locations across three years (Chapter 4). Unique haplotypes were observed for almost every isolate examined. The high level of diversity observed is similar to that reported in an earlier study of *M. nivale* from three turfgrass hosts, wherein 96 out of 100 isolates tested displayed a unique genotype when amplified by RAPD (Mahuku et al. 1998). The ISSR and SSR primers identified in Chapter 4
could be used to assess genetic diversity among isolates of *M. nivale* and *M. majus* collected from a variety of grass and cereal hosts. This investigation could help to clarify whether fungal isolates may be capable of attacking several different host plants in the field. Outside of identifying differences between fungal populations on different host plant species, the application of ISSR and SSR to isolates collected from a single economically important host, such as *T. aestivum*, may be valuable.

A priority for future research pertaining to the genomic differences observed is the sequencing of at least one additional turf isolate of *M. nivale*. This would help to clarify whether some of the differences observed between the wheat and turf isolates are truly due to host-specific variation. Alternatively, it may be possible that turf-derived isolates as a group may be more diverse than wheat-derived isolates; although this direct comparison was not assessed within the thesis, the ISSR and SSR work confirmed that turfgrass isolates of *M. nivale* are indeed genetically diverse. The investigation of *M. nivale* and *M. majus* isolates from other cereals, such as oats and barley would clarify not only the relationship of these pathogen populations to those examined, but may also help to demonstrate whether differing hosts really do represent a barrier to genetic flow among these species.

The yearly emergence of distinct genotypes observed among *M. nivale* is generally consistent with earlier hypotheses (e.g. (Mahuku et al. 1998)) that sexual reproduction is infrequent but present in this species. To explore the genetic basis for mating in *M. nivale* and *M. majus*, an attempt was made to identify the mating-type genes in those species using PCR primers designed for other filamentous ascomycetes (Chapter 5). These initial experiments failed even after considerable effort, and the MAT1 genes were only identified after the genome sequencing data became available. This difficulty was explained by the poor nucleotide-level
conservation of the MAT1 genes in general. Despite previous reports that *M. majus* is homothallic (e.g. (Maurin et al. 1995)) and that *M. nivale* may be either homothallic or heterothallic (Litschko and Burpee 1987; Maurin et al. 1995), only MAT1-2-1 was detected in the genomes of all six *Microdochium* isolates sequenced. When additional isolates of both species were examined by PCR, the candidate MAT1-2-1 sequence was amplified in every isolate tested, regardless of host plant or geographic origin. Furthermore, the sequences were detected between the APN2 and SLA2 genes, which have been reported to flank the MAT1 genes among the Sordariomycetes (Butler 2007). This synteny suggests that the sequence found truly was MAT1-2-1. A candidate sequence for MAT1-1-1 was not detected among any of the *Microdochium* isolates nor among the genomes of any of the other Xylariales examined. Despite possessing only a single mating type, unpaired isolates of *M. majus* produced asci and ascospores when incubated on wheat straw, demonstrating apparently homothallic reproduction without one of the two genes reported to be essential for this process (Butler 2007). Perithecia were observed in some *M. nivale* pairings, but no ascospores were observed; although this result is in general agreement with reports that *M. nivale* produces perithecia less frequently than *M. majus* (Lees et al. 1995), the difference was not readily explained by the apparent presence of only the MAT1-2-1 gene among more than 90 *M. nivale* isolates tested.

These observations in addition to the lack of a candidate MAT1-1-1 homolog from any of the other members of the order Xylariales led to the hypothesis that a yet-undetected gene may be responsible for the control of mating among this order. Other species that are not easily classified as homothallic, heterothallic, or pseudohomothallic based on observations of their mating behaviour or genetic information have been identified among other genera of Sordariomycetes (Chen et al. 2002; Lin and Heitman 2007; Menat et al. 2012; Rodriguez-Guerra
et al. 2005; Vaillancourt et al. 2000) and corroborate the need for a more thorough investigation and analysis of the control of mating among the Ascomycota.

There are several ways to gather further information about the control of mating among Microdochium and other members of the order Xylariales. First, among M. nivale, it would be valuable to perform additional mating crosses in the lab, especially from a wider variety of sources, in order to more thoroughly assess the conditions under which M. nivale produces ascospores (Lees et al. 1995; Parry et al. 1995). This type of investigation would also prove useful among other species within the order Xylariales, for which there is very limited information available regarding the mating style of most species as homothallic, heterothallic, or unclassified. In cases where mating has been observed and can be reliably induced in the lab, a technique like RNA-Seq could be used to reveal which genes are expressed immediately prior to and during the formation of sexual structures and spores. This investigation could reveal candidate genes that may be responsible for the initiation of sexual reproduction in these species, and may reveal the presence of a MAT1-1-1-like sequence that was overlooked in these investigations.

The numerous attempts to identify mating-type genes among Microdochium and other members of the order Xylariales demonstrate the power of having a genome sequence available. Whereas various PCR-based strategies were applied to detect MAT-1 sequences among Microdochium for approximately one year, they were all ultimately unsuccessful and failed to provide any useful results despite consuming a considerable amount of effort and resources. Alternatively, even the very rough draft genome sequence first obtained for M. majus 99049 provided a MAT1-2-1 sequence within literally minutes following assembly. Moving forward, sequencing the genome of an organism of interest, especially when genome sequences of a
closely-related species are unavailable, should be considered for finding genes that are not highly conserved prior to undertaking a project that will include PCR of any sort.

Outside of mating, transposable elements may also present an important source of genetic diversity among fungi, especially pathogens (e.g. (Amyotte et al. 2012; Dean et al. 2005; Hatta et al. 2002; Thon et al. 2006)). Transposable elements were detected among the Microdochium genomes examined (Chapter 3). Transposons in fungi are subject to degradation by a mechanism known as repeat-induced point mutation (RIP) (Selker et al. 1987), which allows the rapid accumulation of mutations within highly repetitive DNA sequences, such as transposons, presumably to prevent them from moving throughout the genome and disrupting functional genes. Interestingly, because TEs are often physically associated with pathogenicity-related genes (such as effectors), this mechanism may encourage the rapid diversification of pathogenicity-related sequences, providing a partial explanation for the high mutation rate of effectors (de Jonge et al. 2011); indeed, the frequency of RIP in non-repetitive regions is correlated with their proximity to repetitive regions (Van de Wouw et al. 2010). In examining the sequenced Microdochium genomes, more than 50% of the putative TEs identified were found to be within 5 kb of a putative pathogenicity-associated gene. A high level of physical proximity between pathogenicity-related genes and highly variable genetic regions, including transposable elements, has been identified among other fungal phytopathogens including Magnaporthe grisea (Dean et al. 2005), L. maculans (Van de Wouw et al. 2010), and Fusarium oxysporum f. sp. lycopersici (Ma et al. 2010).

To assess whether the putative pathogen-host interaction (PHI) genes identified in the genome really do play a role in pathogenicity, RNA-Seq could be used to investigate transcriptional differences in these fungi while they are interacting with a host plant, in
comparison to when they are growing on artificial media. Among the fungal transcripts that are detected at significantly higher or lower levels in the plant-fungal interaction database, RT-PCR could be used to monitor expression level changes of these genes throughout several points in the infection process. Towards this goal, RNA-Seq data have been collected for *M. nivale* and *M. majus* in interaction with *T. aestivum* and *Agrostis stolonifera* (creeping bentgrass). The analysis of those data is ongoing.

Fungal effectors are proteins that interact with and can shut down a plant's immune system (Maffei et al. 2012). When fungal effectors are produced, they may be secreted in either the apoplast or inside plant cells either via haustoria or by co-opting existing plant transporters (Hogenhout et al. 2009); based on the apparent lack of formation of haustoria in the *Microdochium* isolates examined, effector secretion by *M. nivale* and *M. majus* may be in the apoplast. This could be investigated using in situ hybridization to visualize the localization of these enzymes in combination with microscopic investigations of the type already performed.

In this thesis, a variety of techniques were applied to explore differences between the closely-related fungal plant pathogens *Microdochium nivale* and *M. majus*. Traditional phytopathological techniques were combined with tools from modern molecular biology to investigate genetic, genomic, and pathological differences between these two fungal species. Unexpected trends were detected between, but also within these species, and suggest that future work may reveal further differences between populations that are isolated either geographically or on different host species. In addition, the mating type genes found within these species failed to correlate as predicted with the homothallic mode of reproduction that was observed for *M. majus* and that has been reported for *M. nivale*, suggesting that the control of mating in these species, and perhaps in other members of the Xylariales, may involve yet-to-be-identified genes.
relative to other members of the Sordariales. Together, the results in this thesis demonstrate the complexity that remains to be uncovered even within economically important pathogenic species that have been known to science for over 100 years.
7.3 References for Chapter 7


