Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.
Secondary metabolism of the forest pathogen *Dothistroma septosporum*

A thesis presented in the partial fulfilment of the requirements for the degree of Doctor of Philosophy (PhD) in Genetics at Massey University, Manawatu, New Zealand

Ibrahim Kutay Ozturk

2016
ABSTRACT

*Dothistroma septosporum* is a fungus causing the disease Dothistroma needle blight (DNB) on more than 80 pine species in 76 countries, and causes serious economic losses. A secondary metabolite (SM) dothistromin, produced by *D. septosporum*, is a virulence factor required for full disease expression but is not needed for the initial formation of disease lesions. Unlike the majority of fungal SMs whose biosynthetic enzyme genes are arranged in a gene cluster, dothistromin genes are dispersed in a fragmented arrangement. Therefore, it was of interest whether *D. septosporum* has other SMs that are required in the disease process, as well as having SM genes that are clustered as in other fungi.

Genome sequencing of *D. septosporum* revealed that *D. septosporum* has 11 SM core genes, which is fewer than in closely related species. In this project, gene cluster analyses around the SM core genes were done to assess if there are intact or other fragmented gene clusters. In addition, one of the core SM genes, *DsNps3*, that was highly expressed at an early stage of plant infection, was knocked out and the phenotype of this mutant was analysed. Then, evolutionary selection pressures on the SM core genes were analysed using the SM core gene sequences across 19 *D. septosporum* strains from around the world. Finally, phylogenetic analyses on some of the SM core genes were done to find out if these genes have functionally characterised orthologs.

Analysis of the ten *D. septosporum* SM core genes studied in this project showed that two of them were pseudogenes, and five others had very low expression levels *in planta*. Three of the SM core genes showed high expression levels *in planta*. These three genes, *DsPks1*, *DsPks2* and *DsNps3*, were key genes of interest in this project. But despite the different expression levels, evolutionary selection pressure analyses showed that all of the SM core genes apart from the pseudogenes are under negative selection, suggesting that *D. septosporum* might actively use most of its SMs under certain conditions.
In silico predictions based on the amino acid sequences of the proteins encoded by SM core genes and gene cluster analyses showed that four of the SM core genes are predicted to produce known metabolites. These are melanin (DsPks1), cyclosporin (DsNps1), ferricrocin (DsNps2) and cyclopiazonic acid (DsHps1). Gene cluster analyses revealed that at least three of the D. septosporum SMs might be produced by fragmented gene clusters (DsPks1, DsNps1, DsNps2). This suggested that dothistromin might not be the only fragmented SM gene cluster in D. septosporum.

According to phylogenetic analyses, some of the D. septosporum SM core genes have no orthologs among its class (Dothideomycetes), suggesting some of the D. septosporum SMs may be unique. One such example is the metabolite produced by DsNps3. Comparison of wild type and ΔDsNps3 D. septosporum strains showed that the ΔDsNps3 strain produces fewer spores, less hyphal surface network at an early stage of plant infection, and lower levels of fungal biomass in disease lesions compared to wild type, suggesting that the DsNps3 SM may be a virulence factor. Attempts to identify a metabolite associated with DsNps3, and to knockout another gene of key interest, DsPks2, for functional characterization were unsuccessful.

Further work is required to confirm the gene clusters, characterise the SMs and their roles. However, the findings so far suggest that dothistromin is unlikely to be the only D. septosporum SM that is a virulence factor in since the DsNps3 SM also appears to be involved in virulence. Likewise the fragmented dothistromin cluster may not be the only one in the genome and there may be at least three more fragmented SM gene clusters.
Acknowledgements

First of all, I am immensely grateful to my supervisor Dr. Rosie Bradshaw for her support and encouragement throughout my PhD. Even when none of my experiments were working and I wasn't believing in myself, she kept believing in me and supported me in every way possible. I don't believe there is a word in any language to express my level of gratitude. I could not be where I am now without her support.

I also want to thank my co-supervisors Dr. Rebecca McDougal and Dr. Carla Eaton for their valuable suggestions and constructive criticism whenever I asked for assistance.

I also acknowledge Andre Sim, Dr. Sinan Ugur Umu, and especially Dr. Pierre Yves-Dupont for their assistance with bioinformatics tools. Special thanks to Pierre for teaching me some bioinformatics and providing critical readings.

I also want to thank all former and current members of the “Fungal Jungle” lab group. I specially want to thank our lab manager Mrs Carole Flyger for her assistance and rapid help in finding everything I needed. I can't thank Pranav Chettri enough for helping me solve many problems regarding experiments, advising with the experimental setups, and our valuable discussions about my results. He was always willing to help even when he was too busy. I also want to specially thank Melissa Guo for her valuable advice and support. I also want to thank Lukas Hunziker for having patience to set up experiments with me until midnight. I also specially thank Md. Kabir for his valuable suggestions and help even after he graduated. I want to extend my thanks also to Yanfei, Andre and Simren. I feel very lucky to be part of this group.

I sincerely thank Dr. Mark Patchett for his valuable suggestions on the optimization of solvent systems.
I thank my family members at home, Galip, Nimet, Buket, and Elifnur Ozturk as well as Mustafa, Sema, and Halil Yalinkilic for their endless support and faith in me. I always felt their good wishes supporting me, even from thousands of kilometers away. I want to extend my gratitude to Tarcin Yalinkilic for making my family smile even when we were very depressed.

Most importantly, I want to thank my beloved wife Aslinur. I can't thank her enough for her love, support, and sacrifice in this long journey. She always did her best to help me with anything I asked, sacrificed from her precious sleep, always picked me up or brought me outstanding meals anytime I wanted. I have to give my special thanks for her incredible baking skills, which gave me energy and motivation to study long hours. I feel I am the most lucky man in the world for having such a great, understanding, and supportive wife.

I am very grateful to BioProtection Research Centre for funding me for the three years of my project. I would also like to thank the Institute of Fundamental Sciences, IFS, for providing financial assistance after the third year of my PhD.

Finally, I want to thank all good, smiling people of New Zealand. It feels great to be part of such a warm, welcoming community.
Table of Contents

ABSTRACT ................................................................................................................................... i
Acknowledgements ................................................................................................................... iii
Table of Contents ....................................................................................................................... v
List of Tables ................................................................................................................................. x
List of Figures ................................................................................................................................. xi
List of Abbreviations ..................................................................................................................... xiv
1. Introduction ............................................................................................................................... 1
  1.1. Dothistroma needle blight .................................................................................................... 1
  1.2. Dothistroma septosporum ................................................................................................... 4
  1.3. Plant-pathogen interactions ............................................................................................... 5
  1.4. Fungal secondary metabolism ........................................................................................... 9
    1.4.1. Polyketide synthases (PKSs) ..................................................................................... 10
    1.4.2. Nonribosomal peptide synthetases (NRPSs) ............................................................. 14
    1.4.3. Hybrid peptide-polyketide synthases (HPSs) ............................................................ 19
    1.4.4. Dothistroma septosporum secondary metabolism .................................................... 21
      1.4.4.1. Dothistromin ........................................................................................................ 22
    1.4.5. Fungal SM gene clusters ............................................................................................. 24
      1.4.5.1. Dothistromin gene cluster ................................................................................... 25
  1.5. Regulation of secondary metabolism .................................................................................. 30
  1.6. Aims and Objectives ............................................................................................................ 32
2. Materials and Methods ............................................................................................................ 34
  2.1. Biological Materials ............................................................................................................ 34
    2.1.1. Dothistroma septosporum strains ............................................................................. 34
    2.1.2. Escherichia coli strains ............................................................................................. 34
    2.1.3. Plant material .............................................................................................................. 35
  2.2. Growth and maintenance of the biological cultures ......................................................... 35
    2.2.1. Growth and maintenance of E. coli ......................................................................... 35
    2.2.2. Growth, maintenance, and sporulation of D. septosporum ..................................... 35
  2.3. DNA extraction, purification, and quantification ............................................................... 36
    2.3.1. Maxiprep fungal gDNA extraction ........................................................................... 36
2.3.2. Miniprep fungal gDNA extraction ................................................................. 37
2.3.3. Plasmid DNA extraction ............................................................................. 37
2.3.4. Agarose gel electrophoresis ....................................................................... 37
2.3.5. Agarose gel extraction of DNA ................................................................. 38
2.3.6. Nucleic acid quantification ......................................................................... 38
2.4. Restriction endonuclease digestion ............................................................... 39
2.5. E. coli transformation .................................................................................... 39
  2.5.1. Competent cell preparation ....................................................................... 39
  2.5.2. Transformation E. coli Top 10 ................................................................. 40
2.6. Polymerase chain reaction (PCR) ................................................................. 40
  2.6.1. Primer design ........................................................................................... 40
  2.6.2. Standard PCR conditions ........................................................................ 40
  2.6.3. Colony PCR for clone screening ............................................................. 41
2.7. Vector construction ...................................................................................... 41
  2.7.1. Vector construction using GATEWAY ...................................................... 42
  2.7.2. Vector construction using OSCAR .......................................................... 44
2.8. DNA sequencing .......................................................................................... 47
2.9. D. septosporum transformation ...................................................................... 47
2.10. Southern blotting and hybridization ........................................................... 48
  2.10.1. Digoxygenin-11-dUTP (DIG) labeling of probes .................................. 48
  2.10.2. Southern blot .......................................................................................... 49
  2.10.3. Hybridization of DIG labeled probe ...................................................... 50
  2.10.4. Signal detection ...................................................................................... 50
2.11. Secondary metabolite extraction, detection, quantification ....................... 51
  2.11.1. Extraction of metabolites ....................................................................... 51
  2.11.2. Thin Layer Chromatography (TLC) ....................................................... 52
2.12. Phenotypic characterization of WT and ΔDsNps3 D. septosporum ............ 52
  2.12.1. Pathogenicity assay ................................................................................. 52
    2.12.1.1. Surface hyphal network determination ........................................... 53
    2.12.1.2. Biomass estimation ........................................................................... 54
  2.12.2. Sporulation rate analysis ........................................................................ 55
  2.12.3. Radial growth rate analysis .................................................................... 55
2.13. Statistical analyses ...................................................................................... 55
2.14. In silico analyses ..................................................................................................... 56
  2.14.1. Domain analysis .............................................................................................. 56
  2.14.2. Secondary metabolite product prediction ................................................... 56
  2.14.3. Phylogenetic analyses ................................................................................... 56
  2.14.4. Comparison of SM core genes from 19 strains of *D. septosporum* .......... 57
  2.14.5. Gene cluster analyses ................................................................................... 58

3. Results and Discussion .............................................................................................. 59
  3.1. Secondary metabolites in *Dothistroma septosporum* ........................................ 59
    3.1.1. Overview of *Dothistroma septosporum* secondary metabolite genes .. 59
    3.1.2. Expression in planta .................................................................................... 62
    3.1.3. Effects of media conditions on secondary metabolism ............................... 64
  3.2. Nonribosomal peptide synthetases ....................................................................... 74
    3.2.1. NPS1 ............................................................................................................. 75
      3.2.1.1. Confirmation of gene model for *DsNps1* ............................................. 75
      3.2.1.2. Comparison of *DsNps1* from 19 strains of *D. septosporum* .......... 78
      3.2.1.3. Gene cluster analysis ............................................................................. 80
    3.2.2. NPS2 ............................................................................................................. 84
      3.2.2.1. Confirmation of gene model for *DsNps2* ............................................. 84
      3.2.2.2. Comparison of *DsNps2* from 19 strains of *D. septosporum* .......... 86
      3.2.2.3. Gene cluster analysis ............................................................................. 88
    3.2.3. NPS3 ............................................................................................................. 91
      3.2.3.1. Phylogenetic analysis ............................................................................. 91
      3.2.3.2. Confirmation of gene model for *DsNps3* ............................................. 94
      3.2.3.3. Comparison of *DsNps3* from 19 strains of *D. septosporum* .......... 97
      3.2.3.4. Gene cluster analysis ............................................................................. 99
      3.2.3.5. Knockout of *DsNps3* ........................................................................... 102
      3.2.3.6. TLC analysis ......................................................................................... 104
      3.2.3.7. Sporulation and growth rate analyses .................................................. 108
      3.2.3.8. Pathogenicity assay .............................................................................. 109
  3.3. Polyketide synthases ........................................................................................... 114
    3.3.1. PKS1 ............................................................................................................. 114
      3.3.1.1. Phylogenetic analysis ........................................................................... 114
      3.3.1.2. Confirmation of gene model for *DsPks1* ........................................... 117
3.3.1.3. Comparison of DsPks1 from 19 strains of D. septosporum.............117
3.3.1.4. Gene cluster analysis........................................................................120
3.3.2. PKS2 .................................................................................................127
   3.3.2.1. Phylogenetic analysis of DsPks2..................................................127
   3.3.2.2. Confirmation of gene model for DsPks2.......................................129
   3.3.2.3. Comparison of DsPks2 from 19 strains of D. septosporum.........130
   3.3.2.4. Gene cluster analysis.................................................................133
   3.3.2.5. Gene knockout attempts.............................................................135
3.3.3. PKS3 .................................................................................................137
   3.3.3.1. Confirmation of gene model for DsPks3.......................................137
   3.3.3.2. Comparison of DsPks3 from 19 strains of D. septosporum.........139
3.4. Hybrid polyketide-nonribosomal peptide synthetases............................141
   3.4.1. HPS1 ..............................................................................................141
      3.4.1.1. Confirmation of gene model for DsHps1.................................141
      3.4.1.2. Comparison of DsHps1 from 19 strains of D. septosporum....143
      3.4.1.3. Gene cluster analysis...............................................................145
   3.4.2. HPS2 ..............................................................................................147
      3.4.2.1. Confirmation of gene model for DsHps2.................................147
      3.4.2.1. Comparison of DsHps2 from 19 strains of D. septosporum....150
      3.4.2.2. Gene cluster analysis...............................................................152
4. Conclusions and future work......................................................................154
APPENDIX .................................................................................................160
   Appendix 1 – Media ..................................................................................160
   Appendix 2 – Buffers and Solutions..........................................................162
   Appendix 3 – Primers ................................................................................164
   Appendix 4 – Gene model confirmation of DsNps1.................................165
   Appendix 5 – D. septosporum strains used in evolutionary selection pressure analyses .........................................................................................166
   Appendix 6 – Gene model confirmation of DsNps2..................................167
   Appendix 7 – Gene model confirmation of DsNps3..................................168
   Appendix 8 – DsNps3-pOSCAR gene knockout construct.......................169
   Appendix 9 – Southern hybridization confirmation of ΔDsNps3...............170
Appendix 10 – Fluorescence microscopy images of WT and ΔDsNps3 D. septosporum infected needles ........................................................................................................... on CD

Appendix 11 – Fungal to plant biomass ratio estimation in WT and ΔDsNps3 D. septosporum infected lesions ..................................................................................................................... 170

Appendix 12 – Binocular microscopy images of WT and ΔDsNps3 D. septosporum infected lesions ............................................................................................................. on CD

Appendix 13 – Gene model confirmation of DsPks1 ................................................................................................................................. 172

Appendix 14 – Initial phylogenetic analyses of DsPks2 ............................................................................................................................ 173

Appendix 15 – Gene model confirmation of DsPks2 ............................................................................................................................. 176

Appendix 16 – DsPks2-GATEWAY gene knockout construct ................................................................. 177

Appendix 17 – DsPks2 gene knockout attempts ......................................................................................... 178

Appendix 18 – Gene model confirmation of DsPks3 ......................................................................................... 180

Appendix 19 – Gene model confirmation of DsHps1 ......................................................................................... 182

Appendix 20 – Gene model confirmation of DsHps2 ......................................................................................... 183

Appendix 21 – Alignments of predicted SM core proteins with best BlastP matches ................................................................. on CD

Appendix 22 – D. septosporum SM core gene and amino acid sequences ........................................... on CD

Appendix 23 – Whole genome sequences used in the phylogenetic analyses ........................................ 184

References ......................................................................................................................................................... 193
List of Tables

Table 1.1. Functionally characterised fungal polyketides................................................ 13
Table 1.2. Functionally annotated NRPSs and their products. ................................. 19
Table 1.3 Some of the well-known products of PKS-NRPS hybrids. ...................... 20
Table 1.4 Number of key secondary metabolite genes in select Ascomycetes........ 22
Table 1.5. Genes in dothistromin fragmented gene cluster................................. 28
Table 2.1 Strains of D. septosporum used in this project...................................... 34
Table 2.2. BP reaction setup for DsNps3-pOSCAR KO construct preparation........ 46
Table 3.1. Complete list of predicted Dothistroma septosporum NZE10 secondary
metabolite core genes................................................................................................. 60
Table 3.2. Percentage of fungal reads in planta during early, mid, and late stages of
infection................................................................................................................ 62
Table 3.3. Genes within DsNps1 putative gene cluster...................................... 81
Table 3.4. Genes within DsNps2 putative gene cluster...................................... 89
Table 3.5. Genes within DsNps3 putative gene cluster ................................... 100
Table 3.6 Lesion number, size and first lesion appearance in wild type and
\( \Delta \text{DsNps3} \) D. septosporum infected pine needles........................................ 112
Table 3.7. Genes within the DsPks1 putative gene cluster................................ 121
Table 3.8. Candidate genes involved in melanin biosynthesis outside of DsPks1
original gene cluster............................................................................................... 125
Table 3.9. Genes within DsPks2 putative gene cluster....................................... 134
Table 3.10. Genes within DsHps1 - DsDma1 putative gene cluster.................... 145
Table 3.11 Genes within DsHps2 putative gene cluster..................................... 152
Table 4.1. Key findings of this project for D. septosporum................................. 158
List of Figures

Figure 1.1. *Dothistroma* spp. global distribution .......................................................... 2
Figure 1.2. Global climatic suitability of DNB based on 1961-1990 climate normals........ 3
Figure 1.3. Pine needles during DNB disease progress .................................................. 4
Figure 1.4. The Zigzag model of plant pathogen interactions ....................................... 6
Figure 1.5. Secondary metabolites involved in plant pathogenicity .............................. 8
Figure 1.6. Reactions catalysed by PKS catalytic domains .......................................... 11
Figure 1.7. Chemical structures of some fungal polyketides ....................................... 13
Figure 1.8. A comparison of ribosomal and nonribosomal protein synthesis ............ 15
Figure 1.9. Examples of reactions catalyzed by NRPS domains .................................. 18
Figure 1.10. Chemical structures of some nonribosomal peptides ............................. 19
Figure 1.11. Domain organization of a typical HPS ..................................................... 20
Figure 1.12. Chemical structures of some hybrid PKS-NRPS products ..................... 21
Figure 1.13. Chemical structures of aflatoxin B1 and dothistromin ............................ 22
Figure 1.14. Organization and predicted pathway of dothistromin gene cluster in *D. septosporum* .................................................................................................................. 27
Figure 1.15. Organization and expression of dothistromin genes .............................. 29
Figure 2.1. Schematic representation of homologous recombination for gene knockout preparation in fungi ................................................................. 42
Figure 2.2. Overview of GATEWAY vector construction .............................................. 44
Figure 2.3. Overview of OSCAR knockout construct preparation ............................. 45
Figure 3.1. Expression of core SM genes in *D. septosporum* during early, mid, and late stages of infection ........................................................... 63
Figure 3.2. TLC analysis of *D. septosporum* wild type and Δ*LaeA* grown in different conditions ................................................................. 66
Figure 3.3. Comparison of SM gene expression and metabolite band intensity in TLC analysis between *D. septosporum* wild type and Δ*LaeA* strains ............... 68
Figure 3.4. Secondary metabolite production under different media conditions ....... 71
Figure 3.5. Domain structure of DsNps1 and its protein alignment with best BlastP hits ......................................................................................... 77
Figure 3.6. *DsNps1* nucleotide alignment of NZE10 with 18 additional strains ......79
Figure 3.7. Phylogenetic analysis *D. septosporum* whole genome sequences of
NZE10 and 18 additional *D. septosporum* genomes..............................................80
Figure 3.8. *DsNps1* and its predicted gene cluster..................................................81
Figure 3.9. Domain structure of *DsNps2* and its protein alignment with best BlastP
hits..............................................................................................................................85
Figure 3.10. *DsNps2* nucleotide alignment of NZE10 with 18 additional strains.....87
Figure 3.11. Synteny of *DsNps2* and *C. fulvum* Nps2 putative gene clusters and
expression profile of *DsNps2* putative gene cluster *in planta*..............................89
Figure 3.12. Phylogenetic tree and alignment of the peptide sequences of *DsNps3*
and its best FASTA hits..........................................................................................92
Figure 3.13. Phylogenetic tree and alignment of the amino acid sequences of
putative *DsNps3* orthologs................................................................. ..........................94
Figure 3.14. Domain structure of *DsNps3* and its protein alignment with closest
homologs..................................................................................................................97
Figure 3.15. *DsNps3* nucleotide alignment of NZE10 with 18 additional genomes.98
Figure 3.16. Organization and expression of *DsNps3* and its neighbouring genes in
early, mid, and late stages of infection.................................................................100
Figure 3.17 Schematic representation of wild type and KO mutants of *DsNps3*...102
Figure 3.18. PCR-based confirmation of purified colonies for the absence of WT
genotype..................................................................................................................103
Figure 3.19. PCR based confirmation of Δ*DsNps3* candidates for the presence of the
KO genotype.......................................................................................................103
Figure 3.20. Solvent optimization for Nps3 SM detection in TLC analysis..........105
Figure 3.21. TLC picture and band profiles of *D. septosporum* wild type and
Δ*DsNps3* strains.................................................................................................107
Figure 3.22. Sporulation and radial growth rates of *D. septosporum* wild type and
Δ*DsNps3* strains.................................................................................................108
Figure 3.23 Hyphal network on pine needle surface infected with wild type and
Δ*DsNps3* *D. septosporum*....................................................................................110
Figure 3.24. Fungal biomass ratio of Δ*DsNps3*-infected lesions to wild type *D.
septosporum* infected lesions................................................................................111
Figure 3.25. Phylogenetic analysis of *DsPks1* ........................................................116
Figure 3.26. Domain structure of DsPks1 and its protein alignment with closest homologs ........................................................................................................................................... 118
Figure 3.27. DsPks1 nucleotide alignment of NZE10 with 18 additional strains ... 119
Figure 3.28. Synteny of DsPks1 and CfPks1 putative gene clusters and expression profile of DsPks1 putative gene cluster in planta ............................................................... 121
Figure 3.29. Melanin biosynthesis pathway in C. heterostrophus ............................................................... 124
Figure 3.30. Expression of candidate genes in the putative fragmented DsPks1 gene cluster ........................................................................................................................................... 125
Figure 3.31. Candidate genes that may be involved in melanin biosynthesis of putative DsPks1 fragmented gene cluster ............................................................................................................................... 126
Figure 3.32. Phylogenetic analysis of DsPks2 and its best FASTA hits ........................................................................................................................................... 128
Figure 3.33. Amino acid alignment of DsPks2 with its closest FASTA hits ........................................................................................................................................... 129
Figure 3.34. Domain structure of DsPks2 and its protein alignment with closest homologs ........................................................................................................................................... 131
Figure 3.35. DsPks2 nucleotide alignment of NZE10 with 18 additional strains ... 132
Figure 3.36. Organization and expression of DsPks2 and its neighboring genes in early, mid, and late stages of infection ........................................................................................................................................... 134
Figure 3.37. Domain structure of DsPks3 and its protein alignment with closest homologs ........................................................................................................................................... 138
Figure 3.38. DsPks3 nucleotide alignment of NZE10 with 18 additional strains ... 140
Figure 3.39. Domain structure of DsHps1 and its protein alignment with closest homologs ........................................................................................................................................... 142
Figure 3.40. DsHps1 nucleotide alignment of NZE10 with 18 additional strains ... 144
Figure 3.41. DsHps1 - DsDma1 putative gene cluster ........................................................................................................................................... 145
Figure 3.42. Domain structure of DsHps2 and its protein alignment with closest homologs ........................................................................................................................................... 149
Figure 3.43 DsHps2 nucleotide alignment of NZE10 with 18 additional strains ... 151
Figure 3.44. DsHps2 putative gene cluster ........................................................................................................................................... 152
List of Abbreviations

A (domain)  Adenylation
ABC (transporter)  ATP-binding cassette
ACP  Acyl carrier protein
AF  Aflatoxin
aLRT  Approximate likelihood ratio test
AT  Acyltransferase
ATMT  *Agrobacterium tumefaciens* mediated transformation
Avr (gene/protein)  Avirulence
BA  Benzoic acid
bp  Base pairs
C (domain)  Condensation
CM  C-methyltransferase
Ct  Cycle threshold
Cyc (domain)  Heterocyclization
ddH2O  Double distilled water
DH  Dehydratase
DM  Dothistroma medium
Dma  Dimethylallyl tryptophan synthase
DMF  Dimethyl formamide
DMSO  Dimethyl sulfoxide
dN  Non-synonymous mutation number
DNA  Deoxyribonucleic acid
DNB  Dothistroma needle blight
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>dpi</td>
<td>Days post-inoculation</td>
</tr>
<tr>
<td>dS</td>
<td>Synonymous mutation number</td>
</tr>
<tr>
<td>DSM</td>
<td>Dothistroma sporulation medium</td>
</tr>
<tr>
<td>ER</td>
<td>Enoylreductase</td>
</tr>
<tr>
<td>EtAc</td>
<td>Ethyl acetate</td>
</tr>
<tr>
<td>EtBr</td>
<td>Ethidium bromide</td>
</tr>
<tr>
<td>EtOH</td>
<td>Ethanol</td>
</tr>
<tr>
<td>ETS</td>
<td>Effector-triggered susceptibility</td>
</tr>
<tr>
<td>F (domain)</td>
<td>Formylation</td>
</tr>
<tr>
<td>FAD</td>
<td>Flavin adenine dinucleotide</td>
</tr>
<tr>
<td>FAS</td>
<td>Fatty acid synthase</td>
</tr>
<tr>
<td>FDR</td>
<td>False discovery rate</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescent protein</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>Hph</td>
<td>Hygromycin B phosphotransferase</td>
</tr>
<tr>
<td>HPS</td>
<td>Hybrid polyketide - nonribosomal peptide synthetase</td>
</tr>
<tr>
<td>HR</td>
<td>Hypersensitive response</td>
</tr>
<tr>
<td>HR-PKS</td>
<td>Highly-reducing PKS</td>
</tr>
<tr>
<td>HST</td>
<td>Host-specific toxin</td>
</tr>
<tr>
<td>kb</td>
<td>Kilo base pairs</td>
</tr>
<tr>
<td>KO</td>
<td>Knockout mutant</td>
</tr>
<tr>
<td>KR</td>
<td>Ketoreductase</td>
</tr>
<tr>
<td>KS</td>
<td>Keto-synthase</td>
</tr>
<tr>
<td>LAS</td>
<td>Leica application suite</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>LB</td>
<td>Luria broth</td>
</tr>
<tr>
<td>MAFFT</td>
<td>Multiple alignment fast Fourier transform</td>
</tr>
<tr>
<td>MCO</td>
<td>Multicopper oxidase</td>
</tr>
<tr>
<td>MeOH</td>
<td>Methanol</td>
</tr>
<tr>
<td>MFS (transporter)</td>
<td>Major facilitator superfamily</td>
</tr>
<tr>
<td>ML</td>
<td>Maximum likelihood</td>
</tr>
<tr>
<td>MMIC</td>
<td>The Manawatu Microscopy &amp; Imaging Centre</td>
</tr>
<tr>
<td>MS</td>
<td>Mass spectrometry</td>
</tr>
<tr>
<td>NaPDoS</td>
<td>Natural product domain seeker</td>
</tr>
<tr>
<td>NCBI</td>
<td>National center for biotechnology information</td>
</tr>
<tr>
<td>NHST</td>
<td>Non-host-specific toxin</td>
</tr>
<tr>
<td>NM</td>
<td>N-methyltransferase</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>Non-norm</td>
<td>Non-normalized</td>
</tr>
<tr>
<td>Norm.</td>
<td>Normalized</td>
</tr>
<tr>
<td>NR-PKS</td>
<td>Non-reducing PKS</td>
</tr>
<tr>
<td>NRP</td>
<td>Nonribosomal peptide</td>
</tr>
<tr>
<td>NRPS/NPS</td>
<td>Nonribosomal peptide synthetase</td>
</tr>
<tr>
<td>Ox (domain)</td>
<td>Oxidation</td>
</tr>
<tr>
<td>PAMP/MAMP</td>
<td>Pathogen/microbe-associated molecular pattern</td>
</tr>
<tr>
<td>PKS</td>
<td>Polyketide synthase</td>
</tr>
<tr>
<td>PMMG</td>
<td>Pine needle minimal medium + glucose</td>
</tr>
<tr>
<td>PR-PKS</td>
<td>Partially reducing PKS</td>
</tr>
<tr>
<td>PT</td>
<td>Product template</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>PTI</td>
<td>PAMP-triggered immunity</td>
</tr>
<tr>
<td>R (domain)</td>
<td>Reduction</td>
</tr>
<tr>
<td>R (gene/protein)</td>
<td>Resistance</td>
</tr>
<tr>
<td>Rf</td>
<td>Retention factor</td>
</tr>
<tr>
<td>RPMK</td>
<td>Fungal reads per million per kilobase</td>
</tr>
<tr>
<td>SAM</td>
<td>S-adenosyl-L-methionine</td>
</tr>
<tr>
<td>SAT</td>
<td>Starter unit acyl-carrier protein transacylase</td>
</tr>
<tr>
<td>SBSPKS</td>
<td>Structure based sequence analysis of polyketide synthases</td>
</tr>
<tr>
<td>SCD</td>
<td>Scytalone dehydratase</td>
</tr>
<tr>
<td>SLR</td>
<td>Sitewise likelihood-ratio</td>
</tr>
<tr>
<td>SM</td>
<td>Secondary metabolite</td>
</tr>
<tr>
<td>ST</td>
<td>Sterigmatocystin</td>
</tr>
<tr>
<td>T</td>
<td>Thiolation</td>
</tr>
<tr>
<td>T3HN</td>
<td>1,3,8-trihydroxynaphthalene</td>
</tr>
<tr>
<td>T4HN</td>
<td>1,3,6,8-tetrahydroxynaphthalene</td>
</tr>
<tr>
<td>TE</td>
<td>Thioesterase</td>
</tr>
<tr>
<td>TLC</td>
<td>Thin layer chromatography</td>
</tr>
<tr>
<td>wpi</td>
<td>Weeks post-inoculation</td>
</tr>
<tr>
<td>WT</td>
<td>Wild type</td>
</tr>
<tr>
<td>μg</td>
<td>Microgram</td>
</tr>
<tr>
<td>μL</td>
<td>Microliter</td>
</tr>
<tr>
<td>μM</td>
<td>Micromolar</td>
</tr>
</tbody>
</table>
1. Introduction

1.1. Dothistroma needle blight

Dothistroma needle blight (DNB) is one of the most harmful pine diseases in both natural forests and plantations spread throughout the world (Drenkhan et al., 2016). Outbreaks of DNB have been observed since the 1950s in *Pinus radiata* plantations of the Southern Hemisphere, as well as in several regions of North America (Gibson 1972, Gibson 1974). The disease significantly affected large-scale *Pinus radiata* planting projects in India (Bakshi and Singh, 1968), East Africa (Gibson, 1974), and Oregon (Peterson and Harvey, 1976). Within the last 50 years, DNB outbreaks have caused serious damage in East African countries, Chile, New Zealand, Canada, and parts of Europe (Gibson, 1974; Woods et al., 2005; Brown, 2005; Woods et al., 2016). The disease has increased its prevalence in the Northern Hemisphere since the 1990s, especially in Canada and a number of European countries (Drenkhan et al., 2016). A recent study showed that DNB has now been observed in 76 countries (Drenkhan et al., 2016) (Figure 1.1). DNB is still a major problem for pine plantations in the Southern Hemisphere, especially New Zealand (Bulman et al., 2013, Rodas et al., 2016).
The infection rate and distribution pattern of DNB is strongly dependent on host density and climate (Guernier et al., 2004). Climatic conditions have a strong influence on DNB infections because temperature affects pathogen growth and survival in winter, while rainfall provides a good opportunity for dispersal of the pathogen (Watt et al., 2009; Woods et al., 2016). Although the main method for dispersal of the pathogen is considered rain splash, recent studies have shown that the pathogen can spread over five times further than previously anticipated, probably because of transport of the spores via mist or wind (Mullett et al., 2016). Because of its suitable climatic conditions, New Zealand was found to be one of the countries most susceptible to DNB since 1961 (Figure 1.2) (Watt et al., 2009, Bulman et al., 2013).
Initial symptoms of DNB are yellow spots on needles, which then turn into red necrotic bands. Red bands appear because of the accumulation of dothistromin (Bassett et al., 1970), and black fruiting bodies emerge during the late stages of infection (Figure 1.3a). The symptoms of the disease usually involve dead tips but a green base of the needle (Figure 1.3b) (Gadgil, 1967; Shain and Franich, 1981; Bradshaw, 2004).

The mortality rate of DNB was originally thought to be very low when it was the sole type of disease in a tree (Woods et al., 2005). Plants weakened by heavy DNB become susceptible to mortality caused by other agents such as Armillaria spp. (Shaw and Toes, 1977; Woollons and Hayward, 1984). In the last two decades, however, serious DNB outbreaks with high mortality rates have been observed with the epidemics in British Columbia that even killed mature native pine trees in an extraordinary occurrence and there are currently also epidemics in the UK and Columbia (Woods, 2003; Woods et al., 2005; Brown and Webber, 2008; Rodas et al., 2016; Woods et al., 2016).
Figure 1.3. Pine needles during DNB disease progression. a) Stages of DNB in a controlled laboratory infection. During the early stage of disease (3-4 weeks post inoculation/wpi), epiphytic growth and penetration of host tissue take place. At 6-8 wpi (middle stage), early lesions appear indicating the start of host necrosis. During the late stage of infection (10-12 wpi), extended lesions and fruiting bodies erupting from the epidermis are clearly visible. Image reproduced from Kabir et al. (2015). b) D. septosporum infected branch of Pinus mugo. Image reproduced from (OEPP/EPPO, 2008)

DNB has significant effects on the economy of some affected countries. For the New Zealand economy, forestry is one of the key components with $5 billion annual revenue and employing 20,000 people (http://www.ffr.co.nz/new-wood-products-partnership-launched; accessed on 14 April, 2016). The estimated cost of DNB for the New Zealand forest industry is $19.8 m per annum, which increased from $6.1 million in two decades (Watt et al., 2011; New, 1989).

1.2. Dothistroma septosporum

Dothistroma septosporum, an ascomycete belonging to the class Dothideomycetes, is one of two fungal species that cause the disease DNB. The other species, D. pini, has a narrower distribution and is less well studied than D. septosporum (Drenkhan et al., 2016). D. septosporum infects more than 107 Pinaceae host species, of which 93 are Pinus species (Drenkhan et al., 2016) and the others are non-pine hosts such as Cedrus and Abies species (Mullett and Fraser, 2015; Drenkhan et al. 2016). D. septosporum is a hemibiotrophic species closely related to the tomato pathogen Cladosporium fulvum (de Wit et al., 2012).
1.3. Plant-pathogen interactions

Plant disease resistance mechanisms are crucial for the host plant protection from insects, pests, infectious bacteria, viruses, and fungi (Brunner et al., 2012). Plant-pathogen interactions involve two-way communication. On one hand, plants must be able to detect and respond to the pathogen. Since plants lack mobile immune cells and no circulatory system to detect the pathogens, they rely on innate immunity of individual cells and the systemic signals originating from the infection site (Spoel and Dong, 2012, Ausubel, 2005). On the other hand, the pathogen needs to manipulate the host plant biology in order to provide itself a suitable environment for growth and reproduction.

Plant defense responses operate at two levels, which can be represented by a zigzag model (Jones and Dangl, 2006) (Figure 1.4). First, the plant detects pathogen (or microbe)-associated molecular patterns (PAMPs/MAMPs) by plant pattern recognition receptors (PRRs) and activates general defense responses named PAMP-triggered immunity (PTI), which is sufficient to repel most pathogens (Boyd et al., 2013). PTI responses include reactive H$_2$O$_2$ accumulation, callose deposition and production of salicylic acid and ethylene (Lamb and Dixon, 1997; Zipfel and Robatzek, 2010). PAMPs are essential molecules highly conserved around a broad spectrum of pathogen species. The best-studied fungal PAMP is chitin, an essential constituent of fungal and insect cell walls (Monaghan and Zipfel, 2012). It was also discovered that PRRs could sense cell wall fragments or peptides released during wounding or infection, which are named damage-associated molecular patterns (DAMPs) (De Lorenzo et al., 2011).

In order to overcome PTI, a pathogen manipulates the host cellular environment. Both suppression of plant defense responses and changing the host cellular environment are required in order to make it suitable for growth and reproduction of the pathogen. This is achieved by the use of proteins collectively named as effectors (virulence factors) that target host defense pathways and metabolism (Koeck et al., 2011). Whole genome sequencing has enabled bioinformatics approaches to identify candidate effectors. Candidate effector proteins can be
selected from small proteins with signal peptides and, in some cases, short conserved sequences (Haas et al., 2009; Dean et al., 2005; Oberhaensli, et al., 2011). For example, the conserved sequences RXLR and LFLAK define some cytoplasmic effectors in *Phytophthora* spp. (Schornack et al., 2009). As a result of co-evolution with pathogens, plants evolved intracellular or membrane-bound receptors called Resistance (R) proteins that recognize specific effectors (called avirulence proteins (Avr) to achieve effector triggered immunity (ETI) that often leads to a hypersensitive response (HR) (Spoel and Dong, 2012). To overcome ETI, pathogens either mutate the effector that can still perform the same activity but is not recognized by host R proteins, or gain new effectors, and plants evolve new R proteins to detect these new virulence factors (Thomma et al., 2011). This R/Avr gene interaction is named as gene-for-gene resistance (Flor, 1942).

**Figure 1.4. The Zigzag model of plant pathogen interactions.** In phase 1, pathogen recognition occurs through detection of PAMPs such as chitin, leading PAMP-triggered immunity (PTI). In phase 2, pathogen effectors interfere with PTI, causing effector-triggered susceptibility (ETS). In phase 3, R proteins recognize effectors (now called Avr), leading effector-triggered immunity (ETI). In phase 4, effectors either mutate or new effectors are gained to avoid detection by R proteins. This applies selective pressure on plant R genes in order to detect the newly evolved virulence factors. Image reproduced from Jones and Dangl (2006).

Fungal secondary metabolites also play a role in plant pathogenicity as effectors. One group of fungal secondary metabolites are known as host-specific toxins (HSTs) (Collemare and Lebrun, 2011) (Figure 1.5a). HSTs are effectors that induce
toxicity only in the host species in the presence of susceptibility target proteins to this specific HST (Friesen et al., 2008). Resistance to HSTs on host plants is only possible in the absence of the dominant susceptibility target gene. Therefore, HST-host gene interactions are usually termed as “inverse gene-for-gene” systems (Friesen et al., 2007; Fenton et al., 2009).

Some examples of HSTs include the nonribosomal peptide victorin from *Cochliobolus victoriae* (Tada et al., 2005), HC-toxin from *Cochliobolus carbonum* (Walton, 2006), polyketide depudecin from *Alternaria brassicicola* (Wight et al, 2009), and T-toxin from *Cochliobolus heterostrophus* (Inderbitzin et al., 2010).

Another group of fungal secondary metabolites are non-host-specific toxins (NHSTs) (Figure 1.5b). These mycotoxins promote necrosis or chlorosis in a spectrum of plant species, causing part or all of the disease symptoms of the pathogen (Collemare and Lebrun, 2011). Fumonisin, an example of a mycotoxin produced by some *Fusarium* spp., enhances both disease severity and host range of the species that produce it (Williams et al., 2007; Glenn et al, 2008). Other examples of NHSTs include cercosporin from *Cercospora* spp. (Choquer et al., 2005), elsinochrome from *Elsinoe fawcettii* (Liao and Chung, 2008), and sirodesmin from *Leptosphaeria maculans* (Elliott et al., 2007). The studies on fungal SMs have shown that the common plant targets for these effectors include mitochondria, chloroplast, ribosome-mediated protein synthesis, and plasma membrane (Mahoney et al., 2003; Kim et al., 2004; Meiss et al., 2008; Collemare and Lebrun, 2011; Dayan et al., 2008).
Figure 1.5. Secondary metabolites involved in plant pathogenicity. a) Victorin, HC-toxin, depudecin, and T-toxin are HSTs. b) Fumonisin, cercosporin, elsinochrome and sirodesmin are NHSTs. Images adapted from Collemare and Lebrun (2011).
1.4. **Fungal secondary metabolism**

Fungi contain many unique and diverse secondary metabolism biosynthetic pathways. Products of some of these pathways include important antibiotics such as penicillin, toxins such as aflatoxin, or sometimes compounds with both beneficial and detrimental characteristics such as ergot alkaloids. Secondary metabolites can be classified according to their effect on human interests such as plant toxins, pharmaceuticals or growth hormones. However, a chemical classification system classifies these compounds according to their general type based on key enzyme classes involved in their biosynthesis: polyketides, nonribosomal peptides, hybrid polyketide-nonribosomal peptides, terpenes (Collado et al., 2007) and indole alkaloids (Xu et al., 2014). Unlike primary metabolites, secondary metabolites (SMs) are not directly required for the growth of the organism at least under laboratory conditions (Keller et al., 2015). However, in some cases the absence of secondary metabolites can result in long-term impairment of survivability/fecundity of the fungi (Turgeon and Bushley, 2010). Secondary metabolite production usually occurs at the end of exponential phase (Calvo et al., 2002).

Initial systematic studies of fungal secondary metabolites began in 1922 by Harold Raistrick, but identification and practical use of the secondary metabolites only started after the discovery of penicillin by Alexander Fleming in 1929 (Raistrick, 1950). Additional roles of SMs such as immunosuppressants and inhibitors were found later on (Omura, 1992; Umezawa, 1972). Increasing interest in secondary metabolites resulted in identification of thousands of these compounds. A study of 1500 compounds that were isolated and characterised between 1993 and 2001 showed that more than half of them have antitumor, antifungal, or antibacterial activity (Pelaez, 2004).

Even though secondary metabolites have many properties of commercial or medical interest, their biological functions from the fungal perspective have gained less attention. From the fungal perspective, SMs can provide competitive advantages against other organisms (Ramaswamy, 2002; Sim, 2001), work as
metal transporting agents (Haas, 2014), enable symbiosis with other organisms (Rohlfs and Churchill, 2011; Boustie and Grube, 2005), act as sex hormones (Guzmán-de-Peña et al., 1998; Champe et al., 1987) and affect differentiation of the fungi such as sporulation and germination (Schroeder and Johnson, 1995; Kawamura et al., 1999). Some SMs activate sporulation (Calvo et al., 2001; Champe and el-Zayat, 1989; Mazur et al., 1991) and some are pigments that provide resistance against UV light (Kawamura et al., 1999). Germination can either be stimulated or inhibited through use of secondary metabolites (Eaton and Ensign, 1980; Lazaridis, 1980, Charlang et al., 1982, Horowitz et al., 1976). Loss of some secondary metabolites causes reduced growth rate and conidiation (Calvo et al., 2002).

1.4.1. Polyketide synthases (PKSs)

Polyketides are a large diverse group of natural products produced by plants, bacteria, marine organisms, and fungi (Cox, 2007). Polyketides are not only sources of many medicines such as lovastatin, but also virulence factors and toxins of pathogens including aflatoxin and fumonisin (reviewed by Hertweck, 2009). Although studies on polyketides were reported as early as 1893 recent advances in genetic tools enabled rapid accumulation of new information regarding these compounds over the last few decades (Cox, 2007). Thanks to the availability of genome sequences, many novel polyketide compounds such as asperthecin of A. nidulans have also been discovered (Chiang et al., 2013).

Polyketide synthesis is very similar to fatty acid synthesis, as first noted by Birch and Donovan (1953), and their reaction mechanisms have been studied in detail (reviewed by Staunton and Weissmann, 2001; Cox, 2007). Polyketide synthesis involves a conserved general mechanism using iterative elongation of carboxylic acid units such as propionate and acetate (Chooi and Tang, 2012). Both PKS and fatty acid synthases (FAS) condense these carboxylic acid units with malonyl thioesters to form carbon-carbon bond in a Claisen-condensation reaction catalysed by keto-synthase (also called β-ketoacyl-synthase; KS or KAS). The malonyl thioester unit used in this reaction is carried by an acyl carrier protein.
(ACP), and acyl groups are transferred from CoA onto KS and ACP components by an acyl transferase (also called malonyl-acetyl transferase; AT or MAT) domain. These three domains are essential for almost all PKS and FAS proteins (Chooi and Tang, 2012). Some PKSs and FASs use additional domains to further modify the product by reduction with a ketoacyl reductase (KR), dehydratase (DH) or enoyl reductase (ER). Figure 1.6 illustrates reactions catalysed by PKS domains. Even though PKS and FAS have many similarities, the PKS have a few capabilities that FAS lacks (reviewed by Cox, 2007). For example, PKS can control the reduction level during each condensation cycle. In addition, PKS can control starter and extender unit selection as well as the level of chain methylation in each cycle.


Organization and usage of these domains indicates the level of reduction of a PKS product and therefore the type of PKS. A non-reducing PKS (NR-PKS) does not contain any of the additional KR, DH, or ER domains, however, it may contain other functional domains including starter unit ACP transacylase (SAT), product template (PT) or thioesterase (TE) domains (Chooi and Tang, 2012). A highly-reducing PKS (HR-PKS) contains all three KR, DH, and ER reduction domains. A partially reducing PKS (PR-PKS), such as 6-methylsalicylic acid synthase of
*Penicillium patulum*, has only a single reduction domain (Beck et al., 1990; Fujii et al., 1996).

Based on their protein architecture, PKSs are divided into three groups. Most fungal PKSs are type I, very large, multidomain proteins that usually use each domain iteratively (Castoe et al., 2007). Type II PKSs are present in bacteria, and each domain is found in an individual protein instead of multiple domains being grouped in a single large protein as in type I PKSs. Type III PKSs are composed of only KS domains, and are present in plant, fungi and bacteria (Cox, 2007).

Fungal PKSs can catalyse the synthesis of a broad range of polyketides. This can be a simple compound such as orsellinic acid or a complex metabolite like T-toxin. Biosynthesis of polyketides usually requires other enzymes modifying the basic PKS product. Most of the polyketides require activity of multiple proteins encoded by genes usually localized in close proximity to each other, forming a gene cluster. A well-characterised polyketide gene cluster is aflatoxin from *A. parasiticus* (Yu et al., 2004). Some of the genes in aflatoxin gene cluster are predicted to encode reductase, dehydrogenase, P450 monoxygenase and alcohol dehydrogenase enzymes. *D. septosporum* produces a similar compound to aflatoxin, named dothistromin. The dothistromin gene cluster and it’s biosynthetic pathway is explained in detail at Section 1.4.5.1. Figure 1.7 and Table 1.1 show some of the well-known fungal polyketides.
Table 1.1. Functionally characterised fungal polyketides.

<table>
<thead>
<tr>
<th>Organism</th>
<th>PKS name</th>
<th>Polyketide product</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Aspergillus parasiticus</em></td>
<td>PksA</td>
<td>Aflatoxin*</td>
<td>Minto and Townsend (1997)</td>
</tr>
<tr>
<td><em>Aspergillus nidulans</em></td>
<td>PksA</td>
<td>Sterigmatocystin*</td>
<td>Minto and Townsend (1997)</td>
</tr>
<tr>
<td><em>Dothistroma septosporum</em></td>
<td>PksA</td>
<td>Dothistromin*</td>
<td>Bradshaw et. al. (2006)</td>
</tr>
<tr>
<td><em>Cochliobolus heterostrophus</em></td>
<td>PKS18</td>
<td>Melanin*</td>
<td>Eliahu et al. (2007)</td>
</tr>
<tr>
<td><em>Wangiella dermatitidis</em></td>
<td>WdPKS1</td>
<td>Tetrahydroxynaphthalene</td>
<td>Feng et al. (2001)</td>
</tr>
<tr>
<td><em>Monascus purpureus</em></td>
<td>plscT</td>
<td>Citrin</td>
<td>Shimizu et al. (2005)</td>
</tr>
<tr>
<td><em>Penicillium expansum</em></td>
<td>PePks</td>
<td>Patulin*</td>
<td>Sanzani et al. (2012)</td>
</tr>
<tr>
<td><em>Aspergillus terreus</em></td>
<td>atX</td>
<td>6-MSA</td>
<td>Fuji et al. (1996)</td>
</tr>
<tr>
<td><em>Glarea lozoyensis</em></td>
<td>plsc2</td>
<td>6-MSA</td>
<td>Lu et al. (2005)</td>
</tr>
<tr>
<td><em>Aspergillus terreus</em></td>
<td>lovB</td>
<td>Lovastatin</td>
<td>Kennedy et al. (1999)</td>
</tr>
<tr>
<td><em>Penicillium citrinum</em></td>
<td>mkA</td>
<td>Compactin</td>
<td>Abe et al. (2002)</td>
</tr>
<tr>
<td><em>Phoma spp.</em></td>
<td>PhPKS1</td>
<td>Squalestatin tetraetide</td>
<td>Cox et al. (2004)</td>
</tr>
<tr>
<td><em>Cochliobolus heterostrophus</em></td>
<td>Pks1</td>
<td>T-toxin*</td>
<td>Baker et al. (2006)</td>
</tr>
<tr>
<td><em>Fusarium moniliforme</em></td>
<td>ORF3</td>
<td>Fumonisin B1*</td>
<td>Proctor et al. (1999)</td>
</tr>
<tr>
<td><em>Cercospora spp.</em></td>
<td>CTB1</td>
<td>Cercosporin*</td>
<td>Newman et al. (2012)</td>
</tr>
</tbody>
</table>

Table adapted from Cox (2007).
An asterisk indicates polyketides with known importance to plants.

Figure 1.7. Chemical structures of some fungal polyketides.
1.4.2. **Nonribosomal peptide synthetases (NRPSs)**

Nonribosomal peptide synthetases (NRPSs), present in fungi and bacteria, are the core proteins catalyzing biosynthesis of peptides without the use of ribosomes (Turgeon and Bushley, 2010; Strieker, 2010). Many nonribosomal peptides have pharmacological importance and are used for a variety of purposes such as antibiotics, immunosuppressive and cytostatic agents (Schwarzer et al., 2003). Nonribosomal peptide studies are important not only because of the peptides’ current roles in modern medicine, but also because of the possibilities to develop new drugs by using different strategies such as site-directed mutagenesis of substrate binding pockets (Finking and Marahiel, 2004; Williams, 2013).

Nonribosomal peptide synthesis is a well-studied process and its shows some similarities with ribosomal machinery. NRP synthesis requires a minimum of three domains: adenylation (A), peptidyl carrier protein (PCP) (also called thiolation (T)), and condensation (C) (Figure 1.8B). The A domain first activates one amino acid (or hydroxy acid) as amino acyl adenylate, and then enables covalent bonding to the cofactor of the PCP domain called the phosphopantetheine attachment site. Afterwards, the C domain catalyzes the peptide bond formation between two amino acids (Lautru and Challis, 2004). This is why first modules of most NRPSs don’t contain a C domain. Exceptions are named as starter C domains, which acylate the first amino acid with a fatty acid (Rausch et al., 2005). The NRPS modules with domains ordered as C-A-PCP are named Type A, linear NRPS. Type B NRPSs use all their modules more than once, and are named iterative NRPSs. Type C NRPSs, also known as nonlinear NRPSs, don’t follow the usual C-A-PCP domain organization pattern, but instead use some domains more than once (Felnagle et al., 2008). Monomodular NRPSs have one of each A, PCP, and C domains, and multimodular NRPSs have repetitive A, PCP, and C domains (Glinski et al., 2002; Mootz et al., 2002; Bushley and Turgeon, 2010).

The similarities between nonribosomal and ribosomal peptide synthesis can be observed for each step. The function of the NRPS A domain is similar to aa-tRNA synthetase activity, but the NRPS A domain and aa-tRNA synthetase are
structurally unrelated (Conti et al., 1997; May et al., 2002; Stachelhaus et al., 1999). The function of PCP in NRP synthesis is similar to the function of tRNA in ribosomal synthesis. Finally, peptide bond formation done by the ribosome is catalysed by the C-domain of an NRPS. A key difference between nonribosomal and ribosomal peptide synthesis is the lack of proofreading in NRP synthesis (Weber et al., 2000).

**Figure 1.8 A comparison of ribosomal and nonribosomal protein synthesis**

A. Ribosomal machinery for peptide and protein biosynthesis. B. Nonribosomal machinery. The molecules with similar functions are aligned with each other. Image obtained from Finking and Marahiel (2004).

In addition to the three key domains, there are several other domains that can be present in different NRPSs, making the nonribosomal peptides highly variable (Schwarzer et al., 2003). Heterocyclization (Cyc) domains catalyze peptide bond formation as well heterocyclization of cysteine, serine, and threonine residues in a three-step mechanism, therefore increase rigidity of the NRP and protect it against proteolytic degradation (Rausch et al., 2007; Miller and Walsh, 2001; Patel et al., 2003). In some cases, Cyc-domain activity is followed by oxidation or reduction catalyzed by oxidation (Ox) or reduction (R) domains, respectively. These domains
catalyze two-electron transfers on the heterocycles, therefore further modifying the final products (Reimmann et al., 2001; Du et al., 2000). A thioesterase (TE) domain is present in most of the NRPSs, catalyzing the liberation of the product and thereby terminating the synthesis (Schneider and Marahiel, 1998; Sieber and Marahiel, 2003). Protection of the peptide against proteolytic breakdown is provided by methylation of the ends of the peptide by N- or C-methyltransferases (N-Mt, C-Mt) (Gehring et al., 1998; Reimmann et al., 2001; Walsh et al., 2001). Finally, formylation (F) domains modify by adding a formyl group to the N-termini of the nonribosomal peptides (Schoenafinger et al., 2006). Examples of the reactions catalyzed by these domains are shown in Figure 1.9.

Biosynthetic genes encoding for nonribosomal peptide assembly and the biosynthesis of the amino acids are generally physically close to each other, forming a gene cluster. An example is the gliotoxin gene cluster from Aspergillus fumigatus, in which 8 of 12 genes encode enzymes required for biosynthesis and tailoring of monomers required for gliotoxin biosynthesis (Cramer et al., 2006).

Both ribosomal and nonribosomal systems have advantages over each other, making them ideal for their specific functions. With the proofreading mechanism that nonribosomal peptide synthesis lacks, ribosomal peptide synthesis has high precision for the biosynthesis of primary metabolism proteins. On the other hand, nonribosomal peptides are very variable in structure and, as of 8 November 2016, the database of nonribosomal peptides (Norine) contains 1184 peptides (http://bioinfo.lifl.fr/norine/) and genomic databases have thousands more NRPS gene clusters that haven’t been biochemically characterised yet (Walsh and Fischbach, 2010; Caboche et al., 2008). While ribosomal systems are restricted to 20 amino acids as building blocks, there are currently >530 monomers that can be used as substrates for the NRPS assemblies. In addition to diversification in the selection of the monomers, further diversifications are present during the chain assembly and chain termination by the domains mentioned above, thus producing NRPs with different compositions and structures (Du and Lou, 2010). Finally, there are further post-assembly-line tailoring reactions (Walsh et al., 2001) that can be compared with post-translational modifications (PTM) that are done following
ribosomal biosynthesis. However, while PTM is the only diversifying tool for ribosomal machinery, postassembly-line tailoring is one of the many steps for the NRP synthesis mechanism. Therefore, structural variety is higher for the nonribosomal peptides. Figure 1.10 and Table 1.2 show some well-known NRPSs.
Figure 1.9 Examples of reactions catalyzed by NRPS domains. a) Adenylation, b) Condensation, c) Heterocyclization, d) Oxidation, e) Reduction, f) N-methylation, g) C-methylation, h) Formylation. Domains are abbreviated as A: adenylation, PCP: peptidyl carrier protein, C: condensation, Cy: heterocyclization, Ox: oxidation, R: reduction, NMT: N-methylation, CMT: C-methylation, and F: formylation. Images were obtained from Hur et al. (2012).
Table 1.2. Functionally annotated NRPSs and their products.

<table>
<thead>
<tr>
<th>Organism</th>
<th>PKS name</th>
<th>NRP product</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Bacillus brevis</em></td>
<td>tyrocidin A</td>
<td>Tyrocidine</td>
<td>Mootz and Marahiel (1997)</td>
</tr>
<tr>
<td><em>Bacillus brevis</em></td>
<td>GS 1</td>
<td>Gramicidin S</td>
<td>Hori et al. (1989)</td>
</tr>
<tr>
<td><em>Bacillus licheniformis</em></td>
<td>Bacitracin synthetase</td>
<td>Bacitracin</td>
<td>Konz et al. (1997)</td>
</tr>
<tr>
<td><em>Amycolatopsis orientalis</em></td>
<td>CepA, CepB, CepC</td>
<td>Chloroeremomycin</td>
<td>van Wageningen et al. (1998)</td>
</tr>
<tr>
<td><em>Stigmatella aurantiaca</em></td>
<td>MxcG</td>
<td>Myxochelin A</td>
<td>Silakowski et al. (2000)</td>
</tr>
<tr>
<td><em>Bacillus subtilis</em></td>
<td>DhbF</td>
<td>Bacillibactin</td>
<td>May et al. (2001)</td>
</tr>
<tr>
<td><em>Streptomyces acidiscabies</em></td>
<td>TtxA</td>
<td>Thaxtomin A*</td>
<td>Healy et al. (2000)</td>
</tr>
<tr>
<td><em>Penicillium chrysogenum</em></td>
<td>ACVS</td>
<td>ACV</td>
<td>Smith et al. (1990)</td>
</tr>
<tr>
<td><em>Trichoderma virens</em></td>
<td>Peptaibol synthetase</td>
<td>Peptaibol</td>
<td>Wiest et al. (2002)</td>
</tr>
<tr>
<td><em>Cochliobolus victoriae</em></td>
<td>Unknown</td>
<td>Victorin*</td>
<td>Sweat et al. (2008)</td>
</tr>
<tr>
<td><em>Aspergillus nidulans</em></td>
<td>SidC</td>
<td>Ferricrocin*</td>
<td>Eisendle et al. (2003)</td>
</tr>
<tr>
<td><em>Dickeya didantii</em></td>
<td>IndC</td>
<td>Indigodin</td>
<td>Reverchon et al. (2002)</td>
</tr>
</tbody>
</table>

An asterisk indicates NRPs with known importance to plants.

In addition to polyketides and nonribosomal peptides, some SMs, named PKS-NRPS hybrids, are produced by combined action of both PKS and NRPS biosynthetic pathways (Du et al., 2003). PKS-NRPS hybrids are usually found in fungi and involve a PKS followed by a single NRPS, thereby connecting a polyketide with an amino acid chain (reviewed by Fisch, 2013). PKS-NRPS hybrids are a large and highly variable group of complex structures and include many mycotoxins and...
virulence factors (Wieman et al., 2010; Collemare et al., 2008). Aside from their roles in the producing organism, PKS-NRPS hybrids have important pharmaceutical potential due to their ability to synthesize compounds which are hard to produce by chemical synthesis (Boettger and Hertweck, 2013). Figure 1.11 illustrates the typical domain structure of a PKS-NRPS hybrid.

![Figure 1.11. Domain organization of a typical HPS.](image)


Some well-known products of fungal hybrid PKS-NRPSs are shown in Table 1.3 and Figure 1.12.

<table>
<thead>
<tr>
<th>Organism</th>
<th>HPS name</th>
<th>Hybrid PKS-NRPS product</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Aspergillus nidulans</em></td>
<td>ApdA</td>
<td>Aspyridone</td>
<td>Xu et al. (2010)</td>
</tr>
<tr>
<td><em>Penicillium expansum</em></td>
<td>CheA</td>
<td>Chaetoglobosin A*</td>
<td>Schümann and Hertweck (2007)</td>
</tr>
<tr>
<td><em>Aspergillus</em> clavatus</td>
<td>CcsA</td>
<td>Cyclochalasin E</td>
<td>Scherlach et al. (2010)</td>
</tr>
<tr>
<td><em>Fusarium</em> heterosporum</td>
<td>EqiS</td>
<td>Equisetin</td>
<td>Sims et al. (2005)</td>
</tr>
<tr>
<td><em>Aspergillus</em> terreus</td>
<td>ATEG_00325</td>
<td>Isolavipucine</td>
<td>Gressler et al. (2011)</td>
</tr>
<tr>
<td><em>Metarhizium</em> robertsii</td>
<td>NGS1</td>
<td>NG-391</td>
<td>Hayashi et al. (2002)</td>
</tr>
<tr>
<td><em>Aspergillus</em> fumigatus</td>
<td>PsoA</td>
<td>Pseurotin A</td>
<td>Komagata et al. (1996)</td>
</tr>
<tr>
<td><em>Beauveria</em> bassiana</td>
<td>TenS</td>
<td>Tenellin</td>
<td>Eley et al. (2007)</td>
</tr>
<tr>
<td><em>Beauveria</em> bassiana</td>
<td>DmbS</td>
<td>Desmethyl-bassianin</td>
<td>Eley et al. (2007)</td>
</tr>
<tr>
<td><em>Xylaria</em> spp.</td>
<td>PKS3</td>
<td>Xyrrolin</td>
<td>Phonghanpot et al. (2012)</td>
</tr>
<tr>
<td><em>Fusarium</em> spp.</td>
<td>fusA</td>
<td>Fusarin C</td>
<td>Song et al. (2004)</td>
</tr>
</tbody>
</table>

An asterisk indicates hybrids with known importance to plants.
1.4.4. *Dothistroma septosporum* secondary metabolism

It is proposed that biotrophy in fungi is associated with fewer numbers of SM biosynthetic pathways compared to those associated with necrotrophic or saprophytic lifestyles (Spanu et al., 2010). However, a study by de Wit et al. (2012) has shown the hemibiotroph *D. septosporum* has 11 core secondary metabolite (SM) genes, which is less than that of the closely related biotroph *Cladosporium fulvum*, which has 23 SM genes (Table 1.4). In addition, only three of the key SM core genes are predicted to be homologous, which shows there is a much lower level of similarity between the SM repertoires of the two species compared to genome-level similarities between the two species. The three common genes are predicted to encode enzymes involved in production of melanin, siderophore, and dothistromin (de Wit et al., 2012).

![Chemical structures of some hybrid PKS-NRPS products.](image)
Table 1.4 Number of key secondary metabolite genes in select Ascomycetes.

<table>
<thead>
<tr>
<th>Fungal species</th>
<th>Lifestyle</th>
<th>SM number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dothistroma septosporum</td>
<td>Hemibiotroph</td>
<td>11</td>
</tr>
<tr>
<td>Cladosporium fulvum</td>
<td>Biotroph</td>
<td>23</td>
</tr>
<tr>
<td>Zymoseptoria tritici</td>
<td>Hemibiotroph</td>
<td>20</td>
</tr>
<tr>
<td>Parastagonospora nodorum</td>
<td>Necrotroph</td>
<td>38</td>
</tr>
<tr>
<td>Magnaporthe oryzae</td>
<td>Hemibiotroph</td>
<td>46</td>
</tr>
<tr>
<td>Fusarium graminearum</td>
<td>Necrotroph</td>
<td>30</td>
</tr>
<tr>
<td>Aspergillus nidulans</td>
<td>Saprophyte</td>
<td>45</td>
</tr>
<tr>
<td>Neurospora crassa</td>
<td>Saprophyte</td>
<td>12</td>
</tr>
</tbody>
</table>

Table adapted from de Wit et al. (2012).

1.4.4.1. Dothistromin

Among Dothistroma septosporum secondary metabolites, the polyketide dothistromin is the only one studied in detail (Bassett et al., 1970; Shain and Franich, 1981; Bradshaw, 2004; Schwelm et al., 2009; Schwelm and Bradshaw, 2010; Chettri et al., 2013; Bradshaw et al., 2013; Kabir et al., 2015). When dothistromin was injected into pine needles, appearance of DNB symptoms such as red bands and necrotic regions on the needles was reported (Shain and Franich, 1981). The molecular structure of dothistromin is very similar to an intermediate substrate in synthesis of sterigmatocystin and aflatoxins produced by some species of Aspergillus (Shaw et al., 1978). Using a $^{13}$C labelling method, it was shown that the bistetrahydrofurano side-chain labelling during the dothistromin biosynthesis was identical to that of aflatoxin B$_1$ and sterigmatocystin biosynthesis pathways in Aspergillus spp. (Shaw et al., 1978), suggesting common steps in the biosynthesis of these compounds.

Figure 1.13. Chemical structures of aflatoxin B$_1$ and dothistromin. Image reproduced from Bradshaw (2002).
For a long time, it was questioned whether dothistromin is a pathogenicity factor (obligatory for disease occurrence) or virulence factor (increasing disease severity) (Shaner et al., 1992). Injection of dothistromin into pine needles evoked a strong host defence response and accumulation of benzoic acid (BA), an antimicrobial agent synthesized by plants that is known to inhibit *D. septosporum*, in the plant cells surrounding the lesions (Franich, 1986). Most of the dothistromin was degraded to oxalic acid and CO₂ in the plant tissues after 24 hours, possibly due to the presence of the light, since it has been shown that light affects the breakdown of dothistromin outside of plant tissue (Harvey et al., 1976). The continuous expansion of the necrotic lesions on the pine needles after the majority of injected dothistromin was presumably degraded could have been because of the BA accumulation, since BA is also toxic to plant cells in high concentrations (Franich et al., 1986; Bradshaw, 2004). Exposure of cell suspension cultures of *P. radiata* to *D. septosporum* cell wall fragments caused host defense reactions such as phenolic compound biosynthesis, showing host defense responses could also be activated in the absence of dothistromin (Hotter, 1997).

Generation of dothistromin-deficient *D. septosporum* mutants (Bradshaw et al., 2006) finally enabled the determination of whether dothistromin is a pathogenicity or virulence factor. Schwelm et al. (2009) showed that dothistromin-deficient *D. septosporum* strains were able to complete the disease cycle on the pine needles, confirming that dothistromin is not a pathogenicity factor and led to the idea that that dothistromin may be produced for providing competitive benefits over other microorganisms. However, it was later shown that *P. radiata* needles infected with dothistromin-deficient *D. septosporum* strains had smaller lesions, less sporulation of the pathogen and lower levels of mesophyll colonization than those infected with wild type *D. septosporum*, suggesting that dothistromin production is correlated with increased host colonization and necrosis inducing ability of *D. septosporum*, and showing that dothistromin is a virulence factor (Kabir et al., 2015). Nevertheless, the presence of small lesions on plants infected with dothistromin-deficient *D. septosporum* led to the question of whether another compound such as another SM or proteinaceous effector is required to initiate lesion formation in DNB (Kabir et al., 2015).
Early lesions of needles infected with *D. septosporum* have dark green regions outside of the lesion, termed “green islands” (Kabir et al., 2015). Green islands are regions where the pathogen slows down senescence and maintains photosynthetic activity of the host. This term was first associated with biotrophic organisms, but they were later seen in plants infected with hemibiotrophic or necrotrophic pathogens (Walters et al., 2008). Even though green islands were observed on *D. septosporum* infected pine needles; they were lost or reduced as DNB progressed, which led to the hypothesis that dothistromin may be targeting chloroplasts or chlorophyll (Kabir et al., 2015). The necrotic areas caused by dothistromin deficient mutants had more chlorophyll compared to necrotic areas caused by wild type, suggesting that dothistromin potentially targets chloroplast or chlorophyll (Kabir et al., 2015). Further work needs to be conducted to identify the specific interaction of dothistromin with plant organelles.

Dothistromin is toxic towards plants other than pine, as well as to microorganisms and animal cells (reviewed by Bradshaw, 2004). Dothistromin is known to be mutagenic, clastogenic, and cytotoxic to animal cells, therefore concerns for human health, especially for workers in the forest industry, was brought to discussion (Elliott et al., 1989; Stoessl et al., 1990). Moreover, it has structural similarity to one of the strongest natural carcinogenic molecules, aflatoxin B1 (reviewed by Eaton and Gallagher, 1994). Comprehensive research on the toxicity of dothistromin showed that dothistromin can cause growth inhibition, mutagenicity, chromosomal abnormalities and cell lysis on different types of cells on variety of organisms and mammalian cells (Elliott et al., 1989). Although dothistromin demonstrated carcinogenic properties, its severity level was found to be less than that of aflatoxin B1 (Elliott et al., 1989). Currently measures are taken to prevent exposure to forest workers (Bulman et al., 2004).

1.4.5. **Fungal SM gene clusters**

As mentioned in Section 1.4, there are many types of fungal secondary metabolites with very diverse functions. PKS, NRPS, or HPS are named as core enzymes and required for many of the secondary metabolism pathways apart from terpenes and
indole alkaloids (Brakhage et al., 2009; Hertweck, 2009; Strieker et al., 2010; Crawford and Townsend, 2010). However, most fungal secondary metabolite pathways require use of multiple enzymes together, generally located in close proximity with each other. Products of the genes neighboring the core genes can have varying roles such as enzymes required for a multi-step biosynthetic pathway, transporters, or regulatory elements (reviewed by Keller et al., 2005). Studies on secondary metabolite gene clusters have shown that the genes within each cluster are generally co-expressed (Keller and Hohn, 1997). One well-studied example is the sterigmatocystin biosynthesis gene cluster that contains 25 genes spanning a 60-kb region in \textit{A. nidulans}, all of which are co-expressed under inducing conditions (Brown et al., 1996). However, even though some automated methods to predict gene clusters rely on the co-expression of these genes (Umemura et al., 2015; Andersen et al., 2013), it has been questioned whether or not this is always the case (Turgeon and Bushley, 2010). Indeed, genes for the well-studied SM dothistromin shows that there may be exceptions.

1.4.5.1. Dothistromin gene cluster

Because dothistromin is structurally similar to versicolorin B, an intermediate in sterigmatocystin and aflatoxin biosynthesis, the first candidate genes for dothistromin biosynthesis were determined based on the aflatoxin genes using degenerate PCR primers and hybridization probes (Bradshaw et al., 2006; Zhang et al., 2007). The complete set of dothistromin genes was not found until the \textit{D. septosporum} genome sequence was available several years later, and this revealed that the dothistromin putative gene cluster was fragmented into 6 separate regions in chromosome 12, unlike the tight gene clusters of sterigmatocystin and aflatoxin (Chettri et al., 2013). Similar fragmented dothistromin gene clusters were also found in \textit{D. septosporum} close relatives \textit{Passalora arachidicola} (Zhang et al., 2010) and \textit{C. fulvum} (de Wit et al., 2012). It is not yet known why dothistromin gene clusters are fragmented in these species. One possibility is that fragmentation may allow genes to avoid localized chromatin-level or telomere-associated regulation by positioning away from the telomeres (Palmer and Keller, 2010; Shaaban et al., 2010; Bradshaw et al., 2013). Alternatively, a fragmented gene
cluster may enable pathway diversification by recruitment of genes in different fragments. For example, DotB and DotC of the *D. septosporum* dothistromin fragmented gene cluster have no orthologs in sterigmatocystin and aflatoxin gene clusters but are clustered together with Ver1, the ortholog of a key aflatoxin biosynthetic gene (Bradshaw et al., 2013). Dothistromin is not the only SM with a fragmented gene cluster, as aflatrem gene clusters of *Aspergillus flavus* and *A. oryzae* (Nicholson et al., 2009) and the meroterpenoid gene cluster of *A. nidulans* (Lo et al., 2012) are each distributed in two chromosomes.

The dothistromin biosynthesis pathway is quite well-characterised. Figure 1.14 shows the chromosomal organization of the dothistromin fragmented gene cluster and its proposed biosynthesis pathway (Chettri et al., 2013). Many of the genes in the dothistromin fragmented gene cluster have been functionally characterised by loss of function gene knockout mutants. The first dothistromin gene to be characterised was DsVer1 (formerly known as DsDotA), encoding a ketoreductase (Bradshaw et al., 2002). The core polyketide synthase gene of dothistromin biosynthesis, DsPksA, was subsequently shown to be required for dothistromin biosynthesis (Bradshaw et al., 2006). When ΔDsPksA mutants were supplied with norsolorinic acid or versicolorin A, they were able to produce dothistromin, supporting the hypothesis that the dothistromin biosynthesis pathway has similarities to those of sterigmatocystin and aflatoxin (Bradshaw et al., 2006). Shortly after the functional characterization of DsPksA, a gene predicted to encode versicolorin B synthase, DsVbsA was functionally characterised by gene knock out and complementation assays (Zhang et al., 2007). Continued analyses on dothistromin cluster genes enabled functional characterization of genes predicted to encode alcohol dehydrogenase (DsAdhA), fatty acid synthase (DsHexA), peroxidase (DsDotB), NAD(P) reductase (DsOrdB) and regulators of the dothistromin gene cluster (DsAflR and DsAflJ) (Chettri et al., 2013; Chettri et al., 2014; Chettri et al., 2015). Table 1.5 lists genes involved in dothistromin biosynthesis.
Figure 1.14. Organization and predicted pathway of dothistromin gene cluster in *D. septosporum*. A) Six loci (indicated 1 to 6) showing the fragmented gene cluster of dothistromin. Colorless arrows show genes with no predicted function related to dothistromin biosynthesis. Distances between each gene within a locus are shown as bp. Red numbers to the right of the genes indicate the numbers of putative aflR binding sites. *DsNorB* gene was initially thought to be part of the cluster but functional characterization showed no NorB involvement in dothistromin biosynthesis (Chettri et al., 2014). B) Proposed dothistromin biosynthesis pathway up to versicolorin A based on the orthology to aflatoxin biosynthesis genes. C) Putative biosynthetic pathway from versicolorin A to dothistromin. Image reproduced from Chettri et al. (2013).
Table 1.5. Genes in dothistromin fragmented gene cluster.

<table>
<thead>
<tr>
<th>Name</th>
<th>JGI Protein ID</th>
<th>Predicted function</th>
</tr>
</thead>
<tbody>
<tr>
<td>DsVer1</td>
<td>192193</td>
<td>NAD(P) reductase</td>
</tr>
<tr>
<td>DsDotB</td>
<td>75412</td>
<td>Peroxidase</td>
</tr>
<tr>
<td>DsDotC</td>
<td>75413</td>
<td>MFS transporter</td>
</tr>
<tr>
<td>DsPksA</td>
<td>192192</td>
<td>Polyketide synthase</td>
</tr>
<tr>
<td>DsCypX</td>
<td>139960</td>
<td>P450 monoxygenase</td>
</tr>
<tr>
<td>DsAvfA</td>
<td>75546</td>
<td>NAD(P) reductase</td>
</tr>
<tr>
<td>DsMoxY</td>
<td>75547</td>
<td>Flavin-binding monoxygenase</td>
</tr>
<tr>
<td>DsAflR</td>
<td>75566</td>
<td>Regulatory protein</td>
</tr>
<tr>
<td>DsAflJ</td>
<td>57214</td>
<td>Methyltransferase</td>
</tr>
<tr>
<td>DsEst1</td>
<td>75609</td>
<td>Esterase (alpha/beta hydrolase)</td>
</tr>
<tr>
<td>DsOrdB</td>
<td>75648</td>
<td>NAD(P) reductase</td>
</tr>
<tr>
<td>DsAvnA</td>
<td>57312</td>
<td>P450 monoxygenase</td>
</tr>
<tr>
<td>DsHexB</td>
<td>181128</td>
<td>Fatty acid synthase</td>
</tr>
<tr>
<td>DsHexA</td>
<td>66976</td>
<td>Fatty acid synthase</td>
</tr>
<tr>
<td>DsHypC</td>
<td>75655</td>
<td>Anthrone oxidase</td>
</tr>
<tr>
<td>DsVbsA</td>
<td>75656</td>
<td>VerB synthase (cyclase)</td>
</tr>
<tr>
<td>DsNor1</td>
<td>75691</td>
<td>NAD(P) reductase</td>
</tr>
<tr>
<td>DsAdhA</td>
<td>48495</td>
<td>Alcohol dehydrogenase</td>
</tr>
<tr>
<td>DsVerB</td>
<td>75692</td>
<td>Desaturase (P450 monoxygenase)</td>
</tr>
</tbody>
</table>

* [http://genome.jgi.doe.gov/Dotse1/Dotse1.home.html](http://genome.jgi.doe.gov/Dotse1/Dotse1.home.html)

Table adapted from Chettri et al. (2013).

Transcriptomics analysis of *D. septosporum* throughout its infection cycle (Bradshaw et al., 2016) enabled analysis of dothistromin gene expression (Figure 1.15). Even though most of the genes within the fragmented gene cluster share a similar expression pattern, it is seen that some genes such as and *DotB* have different expression patterns. Therefore, even though co-expression of the genes within the putative gene cluster may support the hypothesis that it is part of the gene cluster, the genes with different expression levels should not be ruled out only because they share different expression patterns.
Figure 1.15. Organization and expression of dothistromin genes. Pink: Secondary metabolite core gene (DsPksA). Purple: Fragmented dothistromin cluster genes. Small white arrows: genes without predicted SM functions. Significant differences between early-mid and mid-late stages were marked with “*” and “+”, respectively (p<0.05). The differences between early and late stages were also significant for all cluster genes. Figure adapted from Bradshaw et al. (2016).
1.5. Regulation of secondary metabolism

Secondary metabolism is generally a complex and resource-consuming process (Calvo et al., 2002). Therefore, SM production is under regulation to avoid unnecessary expenditure of energy and resources. SM regulation in filamentous fungi occurs in many different ways.

As explained in Section 1.4.5, fungal SM biosynthesis usually occurs through gene clusters. Many fungal SM gene clusters are regulated by transcription factors encoded by genes that are generally located near the genes involved in the SM biosynthesis. One well-known example of such regulatory transcription factors are the aflR proteins in sterigmatocystin, aflatoxin, and dothistromin biosynthesis (Section 1.4.5.1). AflR proteins are Zn(II)$_2$Cys$_6$ transcription factors that bind to palindromic DNA sequences within promoter regions of the biosynthetic cluster genes (Ehrlich et al., 1999; Fernandes et al., 1998). Elimination of aflR genes caused loss or decrease of sterigmatocystin, aflatoxin, and dothistromin production in the producing organisms (Yu et al., 1996; Chettri et al., 2013).

In addition to regulators within the gene clusters, SM production is also controlled in fungi using global regulators. One of the most commonly known global regulators of fungal secondary metabolism is a gene predicted to encode a methyltransferase named LaeA (loss of aflR expression) (Bok and Keller, 2004). After its first identification in *A. nidulans*, LaeA was studied extensively and its orthologs were found in many fungal species (reviewed by Jain and Keller, 2013). LaeA deletion mutants had decreased ability to produce many well-known fungal SMs such as sterigmatocystin, aflatoxin and gliotoxin (Bok and Keller, 2004; Amaike and Keller, 2009; Bok et al., 2005). LaeA is thought to regulate SM biosynthesis using chromatin-level modification (Strauss and Reyes-Dominguez, 2011). Another global regulator of fungal SMs is VeA (velvet). In *A. nidulans* VeA regulates SM gene clusters in response to light (Purschwitz et al., 2008) and interacts with other regulatory proteins such as LaeA (Sarikaya Bayram et al., 2010). Studies on *D. septosporum* showed that both DsLaeA and DsVeA are regulators of *D. septosporum* SMs (Chettri and Bradshaw, 2016; Chettri et al.,
However, in contrast to their functionally characterised orthologs in *Aspergillus* spp., *DsLaeA* negatively regulates dothistromin biosynthesis (Chettri and Bradshaw, 2016) and *DsVeA* regulates SMs in a light-independent manner (Chettri et al., 2012).

Apart from the use of cluster-specific and global regulatory proteins, SM gene clusters are also regulated by environmental factors. These factors are carbon source, nitrogen source, temperature, pH, light, and salinity, as well as biotic challenges such as competing microorganisms (reviewed by Takahashi et al., 2013). Some environmental changes modify SM production by use of Cys$_2$His$_2$ zinc-finger proteins such as CreA for carbon signalling (Dowzer and Kelly, 1991) or AreA for nitrogen signalling (Hynes, 1975). Modification of environmental conditions is a common method to activate SM gene clusters that are otherwise silent in laboratory conditions, and such a modification was also attempted in this project which will be explained in Section 3.1.3.

The developmental stage of the fungus is also influential in SM production. This way, fungi can produce specific SMs in certain stages of their life cycle when they are most needed. The most common developmental process associated with fungal SM is sporulation (reviewed by Calvo et al., 2002). This is usually achieved using G–protein signalling pathways that regulate both developmental pathways and SM gene clusters (reviewed by Brodhagen and Keller, 2006).
1.6. Aims and Objectives

*D. septosporum* has 11 SM core genes and only one of them, *DsPksA* used in dothistromin biosynthesis, has been studied previously. Dothistromin is a virulence factor in DNB disease and its production involves an unusual fragmented gene cluster. This research focused on the other 10 SM core genes to define the whole set of *D. septosporum* SMs, and to determine if there are other DNB virulence factors or fragmented gene clusters in addition to dothistromin.

**Aim 1:**
Define the complete set of SMs.

**Objectives:**

a) Determine the *D. septosporum* SM profiles by analysing the effects of different media conditions on *D. septosporum* metabolite production.

b) Confirm the SM core gene models that were predicted by the Joint Genome Institute (JGI) by aligning the protein sequences with close homologs and analysing RNA-seq data that were mapped to the gene models.

c) Assess evolutionary selection pressure on the SM core genes based on gene sequences from 19 *D. septosporum* strains.

d) Determine if there are functional orthologs of the SM core genes in other species by phylogenetic analysis.

**Aim 2:**
Make predictions about gene clusters associated with the SM core genes.

**Objectives:**

a) Identify the genes that might be involved in SM biosynthesis and in close proximity to the SM core genes, by searching for genes with similar predicted functions in known fungal SM gene clusters.

b) Determine if there are other genes required for biosynthesis of known SMs similar to the *D. septosporum* predicted SMs by comparing putative SM gene
clusters of *D. septosporum* with characterised gene clusters for the known SMs in other species.

**Aim 3:**
Identify potential virulence factors involved in Dothistroma needle blight.

**Objectives:**

a) Determine *in planta* expression levels of SM core genes by analysis of existing RNA seq data.

b) Mutate at least one SM core gene by targeted gene knock out.

c) Characterise the phenotype of the SM mutant(s) by comparing wild type and SM mutant(s) radial growth rate, sporulation rate, and virulence on pines.

d) Attempt to chemically characterise candidate virulence factor SMs.
2. Materials and Methods

2.1. Biological Materials

2.1.1. Dothistroma septosporum strains

*Dothistroma septosporum* wild type strain NZE10 (CBS128990) was isolated from an infected *Pinus radiata* needle in New Zealand, and is the strain that was sequenced by the Joint Genome Institute ([http://genome.jgi.doe.gov/Dotse1/Dotse1.home.html](http://genome.jgi.doe.gov/Dotse1/Dotse1.home.html)). Other *D. septosporum* strains used and developed in this project are shown in Table 2.1.

Table 2.1 Strains of *D. septosporum* used in this project.

<table>
<thead>
<tr>
<th>Fungal strain</th>
<th>Plasmid</th>
<th>Genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>NZE10</td>
<td>-</td>
<td>Wild type</td>
<td>Barron (2006)</td>
</tr>
<tr>
<td>FJT96 (ΔDsVeA K01)</td>
<td>pR297</td>
<td>NZE10/ΔDsVeA::hph</td>
<td>Chettri et al. (2012)</td>
</tr>
<tr>
<td>FJT130 (ΔDsLaeA K02)</td>
<td>pR324</td>
<td>NZE10/ΔDsLaeA::hph</td>
<td>Chettri et al. (2016)</td>
</tr>
<tr>
<td>FJT157 (ΔDsNps3 K01)</td>
<td>pR411</td>
<td>NZE10/ΔDsNps3::hph</td>
<td>This study</td>
</tr>
<tr>
<td>FJT158 (ΔDsNps3 K02)</td>
<td>pR411</td>
<td>NZE10/ΔDsNps3::hph</td>
<td>This study</td>
</tr>
</tbody>
</table>

*hph*: Hygromycin resistance gene (hygromycin B phosphotransferase)

2.1.2. Escherichia coli strains

*E. coli* Top10 (F– mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara leu) 7697 galU galK rpsL (StrR) endA1 nupG fhuA::IS2) (Life Technologies, Carlsbad, CA, USA) was used for plasmid production and maintenance.

*E. coli* One Shot® ccdB Survival™ 2 T1R cells (F– mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araΔ139 Δ(ara-leu) 7697 galU galK rpsL (StrR) endA1 nupG fhuA::IS2) were used for production of Gateway® vectors (Life Technologies, Carlsbad, CA, USA) which contain a *ccdB* gene as a selection marker (Bernard & Couturier, 1992; Loris et al., 1999).
2.1.3. **Plant material**

Less than 1 year old *Pinus radiata* seedlings from families with different *D. septosporum* resistance levels were used for inoculation. Four different clones, named P1 to P4, were supplied from Scion, Rotorua, New Zealand and kept in the Massey University greenhouse.

2.2. **Growth and maintenance of the biological cultures**

2.2.1. **Growth and maintenance of *E. coli***

Cells were grown in Luria Broth (LB) with shaking at 180 rpm or on LB agar (Appendix 1) plates at 37°C overnight.

LB agar plates were sealed with parafilm and stored at 4°C for short-term storage. *E. coli* cultures were mixed with 50% (v/v) sterile glycerol to a final concentration of 15% (v/v) and stored at -80°C for long term storage.

2.2.2. **Growth, maintenance, and sporulation of *D. septosporum***

Inoculum for growth and sporulation of *D. septosporum* NZE10 was obtained by cutting the edge of an actively growing colony of *D. septosporum* in a rich medium named Dothistroma medium (DM) (Appendix 1), then grinding the mycelium using a sterile micro pestle in a microcentrifuge tube with 600 μL sterile MilliQ water. Then 200 μL of that suspension was spread onto DM, Dothistroma sporulation medium (DSM) (Appendix 1) or Pine needle minimal medium + Glucose (PMMG) (Appendix 1) plates or inoculated into 25 mL of DM or PMMG broth in 125 mL flasks, followed by incubation at 22°C for 7-15 days. Subculturing onto fresh DM plates was performed every 3 months for maintenance of the *D. septosporum* strains. For longer-term storage, sections of mycelia obtained by cutting the edge of actively growing colonies were ground up using sterile micropestles, and stored in 50% (v/v) glycerol at -80°C.
*D. septosporum* spores were obtained by preparing inoculum as above, then 200 μL of that suspension was spread onto DSM or PMMG plates and incubated at 22°C for 7 days. Approximately 2 mL sterile, double distilled water (ddH2O) was added to the plates, spread with a sterile glass spreader and incubated for 10 minutes to facilitate spore release into the water. The spore suspension was then transferred to a sterile microcentrifuge tube.

For determination of the effects of different media conditions on *D. septosporum* metabolite profiles, variations of PMMG were produced with 3% glucose, 10 μM FeSO47H2O, and NH4NO₃, as in Appendix 1 except: (a) 1% glucose, (b) no FeSO47H2O (c) (NH₄)₂SO₄ as sole nitrogen source (d) KNO₃ as sole nitrogen source.

### 2.3. DNA extraction, purification, and quantification

#### 2.3.1. Maxiprep fungal gDNA extraction

Genomic DNA from *D. septosporum* cultures grown in 5 mL DM in 25 mL flasks for 15 days was extracted using a protocol adapted from a Genomic DNA Mini Kit (Plant) (Geneaid, New Taipei City, Taiwan). First, the mycelia were harvested using a sterile funnel and nappy liners. Harvested mycelia were freeze-dried overnight and powdered in liquid nitrogen using a mortar and pestle, then transferred to microcentrifuge tubes. Afterwards, 400 μL of GP1 (Geneaid) buffer and 5 μL RNase A were added to the samples and vortexed. The samples were then incubated at 65°C for 10 minutes with vortexing every 3 minutes, followed by addition of 100 μL GP2 (Geneaid) buffer, vortexing and incubating in ice for 3 minutes. The samples were then transferred to filter columns in 2 mL collection tubes. The filter columns were centrifuged for 1 minute at 1000× g, and each eluate was mixed with 1.5 volumes of GP3 buffer containing isopropanol in a microcentrifuge tube. The samples were then vortexed immediately for 5 seconds and placed in GD columns in 2 mL collection tubes, centrifuged and the eluates were discarded. GD columns were washed first with 400 μL W1 buffer and then with 600 μL wash buffer containing ethanol, by centrifuging at 16,000 g for 30 seconds after adding each buffer and discarding the eluates. After washing, the columns were centrifuged for
3 minutes to dry the column matrix. The dry columns were transferred into clean microcentrifuge tubes, 100 μL elution buffer pre-heated to 65°C was added and the tubes were incubated for 5 minutes at room temperature. Finally, purified DNA was obtained by centrifuging the columns at 16,000 g for 30 seconds.

### 2.3.2. Miniprep fungal gDNA extraction

Miniprep gDNA extraction was performed for rapid collection of gDNA from numerous colonies to screen for mutant candidates. Genomic DNA from *D. septosporum* was extracted using a CTAB protocol adapted from Zhang et al. (2010). Small sections taken from the edges of growing *D. septosporum* colonies were transferred into microcentrifuge tubes. The sections were then crushed in 600 μL 2% CTAB buffer (Appendix 2) using a sterile micropestle, followed by incubation at 65°C for 10 minutes. After cooling to room temperature, 500 μL chloroform was added, vortexed and centrifuged at 16,200 g for 3 minutes. Then 400 μL of the supernatant was taken and mixed with 600 μL of chilled 96% ethanol (EtOH), inverted several times, and kept at –20°C for 20 minutes. The samples were then centrifuged at 16,200 g for 5 minutes; the pellets were air dried and resuspended in 35 μL TE buffer (Appendix 2).

### 2.3.3. Plasmid DNA extraction

Plasmid from overnight grown (16 hours) *E. coli* cultures (Section 2.2.1) were extracted using a High Pure Plasmid Isolation Kit (Roche) according to the manufacturer’s protocol.

### 2.3.4. Agarose gel electrophoresis

Analysis of PCR products and gDNA was carried out by melting 0.7%-1.5% (w/v) agarose (Gold Bio, St Louis, USA) in 1x TBE buffer (Appendix 2) based on the molecular weight of the DNA sample. Gels were poured and run with 1x TBE buffer in a gel box (Bio-Rad, Hercules, CA, USA).
DNA samples were mixed with 6× loading dye (Appendix 2) and loaded onto the gel along with 1 Kb Plus DNA molecular size marker (Invitrogen, CA, USA). The electrophoresis was carried out at 80 volts for small (50 mL) gels and 40 volts for big (200 mL) gels, until the dye was approximately 2 cm away from the end of the gel. Afterwards, the gel was stained in ethidium bromide (EtBr) solution (1 μg/mL – Appendix 2) for 20 min, rinsed in ddH₂O, visualised with Gel-DocTMXR documentation system (Bio-Rad) and photographed with Image Lab™ software.

2.3.5. **Agarose gel extraction of DNA**

After agarose gel electrophoresis, specific PCR fragments or restriction digestion products were cut out from the gel using a sterile scalpel while visualizing on 365 nm UV transilluminator (Alpha Innotech, Johannesburg, S.A) and transferred to sterile pre-weighed microcentrifuge tubes. The DNA was then recovered from the agarose block using a PureLink™ Quick Gel Extraction Kit (Invitrogen) according to the manufacturer’s instructions.

2.3.6. **Nucleic acid quantification**

Quantification of PCR products and plasmid DNA was performed using a Nanodrop® ND-1000 UV-Vis spectrophotometer with software version 3.1.0 (Nanodrop Technologies Inc, Wilmington) according to the manufacturer’s instructions. Quantification of gDNA samples was performed by using a Hoefer DyNA Quant 200 fluorometer (Amersham Biosciences, UK) according to the manufacturer’s instructions.
2.4. **Restriction endonuclease digestion**

Restriction endonuclease (RE) digestion of DNA was done by mixing 20 units of a selected RE per 1 μg of DNA in a total volume of 50 μL according to the manufacturer’s instructions. The reaction mixtures were incubated in an appropriate temperature overnight to digest the DNA samples. Since the restriction enzymes used in this project (EcoRV and Sacl) could be heat-inactivated, the reactions were stopped by incubating the reaction mixtures at 65°C for 15 minutes.

2.5. **E. coli transformation**

2.5.1. **Competent cell preparation**

The protocol for competent cell preparation was adapted from Chung and Miller (1993). An isolated *E. coli* Top 10 colony (Section 2.2.1) was inoculated into 5 mL LB and incubated at 180 rpm, 37°C overnight. From this culture, 0.5 mL was inoculated into 50 mL LB in a 250 mL flask and shaken at 37°C until the OD$_{600}$ was 0.45-0.6. The culture was chilled on ice for 10 minutes, and equally distributed between two sterile centrifuge tubes. These cells were then centrifuged at 5,320 g for 5 minutes at 4°C in an Eppendorf 5415R microcentrifuge (Eppendorf, Hamburg, Germany). After the supernatant was discarded, the cells were resuspended in 12.5 mL of chilled 0.1 M CaCl$_2$ and left on ice for 10 minutes. The cells were then centrifuged again using the same conditions, but this time resuspended in 1.25 mL of chilled 0.1 M CaCl$_2$. Finally, the cells were left on ice for a minimum of 30 minutes and either used for immediate transformation, or 133 μL of cells were mixed with 67 μL of 50% glycerol/0.1 M CaCl$_2$ and kept at -80°C for long-term storage.
2.5.2. Transformation *E. coli* Top 10

Up to 5 μL of BP or LR reaction products (Sections 2.7.1 and 2.7.2), or approximately 10 ng (<1 μL) of target plasmid suspension was added to 200 μL of competent cells, and the mixture was chilled on ice for 30 minutes. Heat shock was applied by transferring the chilled mixture to 42°C for 2 minutes, followed by immediate chilling in ice for 2 minutes. To the mixture, 1 mL of LB was added and incubated at 37°C for 1 hour for the recovery of bacteria and expression of selection marker encoded by plasmid. From this cell suspension, 50, 100 and 200 μL volumes were spread onto plates of LB-agar containing 50μg/mL of the appropriate antibiotic such as spectinomycin for BP reaction products of OSCAR cloning, and incubated overnight at 37°C. For transformations with low expectancy of positive colonies such as BP or LR reaction products, the cells were pelleted at 10,000 g for 30 seconds. Then half of the supernatant was discarded, the cells resuspended in the remaining supernatant, and equally distributed over 3 plates then incubated at 37°C overnight.

2.6. Polymerase chain reaction (PCR)

2.6.1. Primer design

All primers were designed using Primer3 v2.3.4 (http://primer3.sourceforge.net/releases.php/) and synthesized by Integrated DNA technologies (Coralville, IA, USA). Primers were stored at 200 μM stock concentrations at -20°C and diluted to 5 μM before use. A list of all primers used in this study is in Appendix 3.

2.6.2. Standard PCR conditions

All PCR reactions were carried out in PCR tubes (Axygen, CA, USA) using an Eppendorf Gradient Mastercycler® (Eppendorf). PCR conditions were adjusted for the amplification of single specific amplicons by changing the annealing temperature according to the primer specificities and following the manufacturer's
specifications for the other parameters. Each PCR reaction was carried out in a 25 μL volume containing: 2.5 μL of 10× PCR buffer without Mg (200 mM Tris-HCl (pH8.4), 500 mM KCl), 1.5 mM MgCl₂, 200 μM dNTP, 0.2 μM of each primer, 1-20 ng of template, and 1 U of Platinum™ Taq DNA Polymerase (Invitrogen). Cycling conditions were 94°C initial denaturation for 5 minutes, then 30 cycles of reactions including 94°C denaturation for 30 seconds, annealing for decided annealing temperature for 30 seconds, and extension at 72°C set according to the amplicon size (1 minute/kb), with a final extension at 72°C for 10 minutes. For PCR reactions required in the construction of vectors, 1 U of Platinum® Taq DNA Polymerase High Fidelity was used (Invitrogen). PCR reactions using high fidelity enzyme were carried out in 50 μL volumes and the same constituents as above except the use of 5 μL 10× High fidelity PCR buffer (600 mM Tris-SO₄ (pH8.9), 180 mM (NH₄)₂SO₄) and 1.5 mM MgSO₄.

2.6.3. Colony PCR for clone screening

For the confirmation of transformed *E. coli* clones, cells were collected from colonies using sterile tips and resuspended in 10 μL sterile water for single colony PCR. In case of large-scale screening, cells from up to 10 colonies were pooled in 50 μL sterile water. The cells were then boiled for 10 minutes and kept on ice for 2 minutes. Afterwards, the cells were centrifuged at 21,100 g for 1 minute and 1 μL of supernatant was used as template for PCR reaction.

2.7. Vector construction

One of the most common methods used in functional characterization of genes is the use of gene knockout (KO) mutants. In fungi, targeted gene replacement is achieved by homologous recombination (Weld et al., 2006). This requires transformation of fungi with a vector containing the same DNA sequences as the flanking regions of the gene of interest for recombination and a selectable marker to select the transformants (Figure 2.1). Two different methods for vector
preparation were used, Multisite GATEWAY® and One Step Construction of
Agrobacterium Recombination-ready plasmids (OSCAR).

Figure 2.1. Schematic representation of homologous recombination for gene
knockout preparation in fungi. Image adapted from
http://www.rasmusfrandsen.dk/technologies.htm

2.7.1. Vector construction using GATEWAY

MultiSite Gateway® Technology is a cloning method based on the bacteriophage
lambda site-specific recombination system (Bushman et al., 1985; Landy, 1989;
Ptashne, 1992). Site-specific recombination between phage and E. coli occurs at
specific sequences named att sites. Specific enzymes are required for these
recombination events, and are provided in the GATEWAY system as BP clonase™
II and LR Clonase™ II Plus enzyme mixes (Landy, 1989; Ptashne, 1992).

An overview of GATEWAY vector construction is presented in Figure 2.2. The most
important step for constructing vectors by MultiSite Gateway® Technology is to
design primers with appropriate \textit{att} sites. Primers were designed so that \textit{attB} sites of the PCR fragments recombined with \textit{attP} sites of pDONR vectors after BP reaction, which lead to production of entry clones. For this project, the \textit{DsPks2 KO} construct was prepared using the GATEWAY method. All the reactions were carried out according to manufacturer’s protocols with minor modifications. Donor vectors were prepared in 5 μL BP reactions.

The three entry clones pDONR™P4P1R, pDONR™221, and pDONRP2R-P3 then contained 5’ flanking region of the target gene, selection marker gene (\textit{hph}), and 3’ flanking region of the target gene, respectively. The donor vector for the selectable marker gene was previously prepared (lab number: pR225). All donor vectors contained a kanamycin resistance gene. The constructs for 5’ and 3’ flanking sites were then transformed into Top10 \textit{E. coli}, selected in LB-kanamycin (100 μg/ml) plates, recombinants confirmed by colony PCR using PKS2-73814-KO5F/ PKS2-73814-KO5R for 5’ flanking donor vector and PKS2-73814-KO3F/ PKS2-73814-KO3R for 3’ flanking donor vector and entry plasmids were extracted from the clones as described in section 2.3.3. A 5 μL LR reaction was then carried out by mixing all three entry clones, a destination vector pDEST™R4-R3 containing a different antibiotic (ampicillin) resistance marker and LR clonase in an overnight reaction. Finally, the construct was transformed as above, selected in LB-Ampicillin (100 μg/ml) and confirmed by PCR and sequencing.
Figure 2.2. Overview of GATEWAY vector construction. Three PCR fragments containing attB sites recombine with pDONR vectors with attP sites to produce entry clones containing attL and attR sites. The three entry clones’ attL and attR sites recombine during LR reaction to produce the knockout construct. Image adapted from Multisite Gateway® Three-Fragment Vector Construction Kit manual, Version G, 8 September 2008.

2.7.2. Vector construction using OSCAR

The One Step Construction of Agrobacterium-Recombination-ready-plasmids (OSCAR) methodology starts with amplification of flanking sites of the gene using primers each containing 5’ extensions with different attB recombination sites, modified from the GATEWAY system. These PCR fragments are then mixed with specifically designed marker and binary vectors along with BP clonase, therefore generating the deletion construct in a single step (Figure 2.3). This deletion
construct is then used to replace the region in between the flanking sites with the hygromycin resistance gene.

Figure 2.3. Overview of OSCAR knockout construct preparation. 5’ and 3’ flanking regions of the target sequence are PCR amplified, followed by BP clonase reaction of these PCR products with binary and selection marker plasmids pOSCAR and pA-Hyg-OSCAR plasmids, respectively. Recombination sites for this system include B1r & B2r for 5’ flank PCR, B3 & B4 for 3’ flank PCR, P2r & P3 for pOSCAR, and P1r and P4 for pA-Hyg-OSCAR. The final product was confirmed by PCR. Image from Paz et al. (2011).

One advantage of the OSCAR method over GATEWAY is the use of binary vectors which enable the use of Agrobacterium tumefaciens mediated transformation (ATMT). In addition, the OSCAR method requires only one cloning reaction (BP), while the GATEWAY method requires both BP and LR reactions. Therefore, the OSCAR method requires less time to prepare a knockout construct. Most importantly, one of the donor vectors for GATEWAY system, p4-P1R, was reported to have more than 2 X 10^4 fold increase in mutation rate compared to other vectors of the system (Kiedrowski, 2011), and this was one of the most common problems faced with GATEWAY in this project.

The protocol for DsNps3-pOSCAR KO construct preparation was adapted from Paz et al. (2011). Primer pairs used in PCR amplification of 5’ and 3’ flanking sites were
NPS3-5FP-attb2R/NPS3-5RP-attb1R and NPS3-3FP-attb4_domain/NPS3-3RP-attb3_domain, respectively (Appendix 3).

PCR reactions were carried out on *D. septosporum* NZE10 gDNA as outlined in Section 2.6, with initial optimization of Tm by gradient PCR with 2°C temperature intervals between 50°C-60°C for both the 5' and 3' flank primer pairs. Afterwards, high-fidelity PCRs were performed for both 5' and 3' flanking sites with 50 μL total volumes, 1 minute 8 seconds extension times, and 25 cycles. The PCR fragments were purified using a High Pure PCR Product Purification Kit (Roche Applied Science, Penzberg, Germany) according to the manufacturer’s protocol and the concentrations were measured by NanoDrop® ND-1000 UV-Vis Spectrophotometer with software version 3.1.0 (Nanodrop Technologies Inc, Wilmington). The OSCAR reaction setup was prepared according to Paz et al. (2011), mixing 60 ng of each vector with 10 ng each PCR fragments as in Table 2.2.

<table>
<thead>
<tr>
<th>Amount (ng)</th>
<th>Amount (μL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pOSCAR (243 ng/μL)</td>
<td>60</td>
</tr>
<tr>
<td>pA-Hyg-OSCAR (179 ng/μL)</td>
<td>60</td>
</tr>
<tr>
<td>5’ flank site PCR product (23 ng/μL)</td>
<td>10</td>
</tr>
<tr>
<td>3’ flank site PCR product (27 ng/μL)</td>
<td>10</td>
</tr>
<tr>
<td>BP Clonase™ II enzyme mix (5x)</td>
<td>1</td>
</tr>
<tr>
<td>TE Buffer</td>
<td>0.84</td>
</tr>
<tr>
<td>Water</td>
<td>1.78</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>5</strong></td>
</tr>
</tbody>
</table>

This reaction mixture was incubated at 25°C overnight, then terminated using 0.5 μL proteinase K (20 μg/μL), incubating at 37°C for 10 minutes. The products were then transformed into *E. coli* Top10 competent cells as outlined in Section 2.5, and approximately half of the colonies (~600 μL) were inoculated and selected in LB-spectinomycin (50 μg/mL). The following day, colony PCR was performed on 10 of the spectinomycin-resistant colonies using NPS3-5FP-attb2R/NPS3-3RP-attb3_domain primers at 58°C annealing temperature, 3 minutes extension time and 25 cycles.
2.8. DNA sequencing

Sequencing reactions of recombinant plasmids and PCR products were carried out at the Massey Genome service (Massey University, Palmerston North) using an ABI PRISM® BigDye® Primer Cycle Sequencing Ready Reaction Kit in an ABI 3730 DNA Analyzer (Applied Biosystems, Foster City, CA). For plasmid sequencing, 300 ng purified plasmid DNA was mixed with 3.2 pmol primer, and filled up to a total volume of 15 μL with sterile double distilled (dd) H2O. PCR product sequencing was done by mixing 2 ng/100 bp of PCR product with 3.2 pmol primers and filling with ddH2O to 15 μL. The sequences were analysed using Geneious version 8.0.3 (Kearse et al., 2012) and National Center for Biotechnology Information (NCBI) BLAST tools (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

2.9. D. septosporum transformation

Protoplasts were prepared based on the protocol of Yelton et al. (1984). In 125 mL flasks, 25 mL liquid DM was inoculated as in Section 2.2.2, and incubated at 22°C for 7 days, shaking at 160 rpm. The mycelia were harvested using a sterile funnel and nappy liners, pooling 3 flasks. The mycelia were washed two times with 30 mL sterile water. The harvested mycelia were then washed with 10 mL OM buffer (Appendix 2), and transferred to a sterile flask. Glucanex 200G (Novozymes, Bagsværd, Denmark) was dissolved in 30 mL OM buffer at 10 mg/mL, then filter-sterilized (0.2 μm Minisart acrodisc), added to the pooled and washed mycelia and incubated at 30°C for 12-16 hours, shaking at 80-100 rpm. Protoplast formation was then confirmed under a Zeiss 47 06 00-9901 light microscope (Zeiss, Germany). The protoplasts were filtered through a nappy liner into sterile 15 mL Corex tubes at 5 mL/tube. ST buffer (2 mL) (Appendix 2) was added slowly to form a clear overlay, and then centrifuged at 3220 g for 7 minutes in A-4-62 rotor (Eppendorf). The protoplasts in the interface were removed carefully using a 200 μL pipette, and transferred into 15 mL Corex tubes (5 mL/tube). Five mL of STC buffer (Appendix 2) was added to wash, mixed gently, and then the samples were centrifuged at 3220 g for 7 minutes. This step was repeated three times, and then
the protoplasts were re-suspended in 0.5 mL STC buffer. The protoplast concentration was estimated using a haemocytometer, then diluted to $1.25 \times 10^8$ protoplasts/mL with STC buffer. Transformation was performed immediately to obtain the highest efficiency using procedures depicted in Vollmer and Yanofsky (1986) and Oliver et al. (1987). Protoplasts (80 μL) were mixed with 20 μL of 40% PEG solution, and 5 μg of $DsPks2$-GATEWAY or $DsNps3$-pOSCAR knockout construct. In addition, the pA-Hyg-OSCAR plasmid, which contains the hygromycin resistance gene and the $TrpC$ promoter, was used as a positive control, and empty protoplasts were used as negative control. The samples were vortexed briefly and left on ice for 30 minutes, then 900 μL of 40% PEG solution was added, mixed and incubated at room temperature for 20 minutes. 100 μL of each sample was mixed with 3.5 mL molten 0.8% Regeneration medium (RG) (Appendix 1) at 48°C and inoculated into 1.5% regeneration media (RG) plates as overlays. The plates were incubated overnight at 22°C, and then 5 mL 0.8% RG overlay containing hygromycin was added to a final antibiotic concentration of 70 μg/mL. The plates were incubated for 3 weeks at 22°C. Genomic DNA extraction was performed as detailed in Section 2.3.2.

2.10. Southern blotting and hybridization

2.10.1. Digoxygenin-11-dUTP (DIG) labeling of probes

DIG-labeled probes were prepared by using a PCR based method (Section 2.6), but with 200 μM of each dATP, dCTP, dGTP, 133 μM dTTP and 67 μM alkali-labile Digoxygenin-11-2′-deoxy-uridine-5′-triphosphate (Roche, Penzberg, Germany) in a 25 μL reaction volume.

Too high a probe concentration may cause significant background noise after hybridization. Therefore, optimum probe concentration was determined by preparing 5-fold serial dilutions of the probe with sterile ddH₂O, spotting 2 μL of each onto a Hybond-N+ membrane (GE Healthcare Ltd), and fixing the probe to membrane by using a UV crosslinker (Cex-800 UV crosslinker, Ultra-Lum Inc). Hybridization and detection procedures were done as described in Sections 2.10.3
and 2.10.4. The probe concentration with clear signal but low background was used for Southern hybridization.

2.10.2. Southern blot

In order to confirm integration site and copy number for DsNps3 and DsPks2 KO mutant candidates, Southern blot analysis was performed. First, gDNA belonging to the KO candidates and wild type D. septosporum were extracted (Section 2.3.1) and digested with restriction enzymes (Section 2.4). Then, the digested gDNA were separated by agarose gel (1%) electrophoresis at 40 V overnight, and visualized as in Section 2.3.4. The gel was then depurinated by gently agitating in Solution 1 (depurination) (Appendix 2) for 15 min. Afterwards, solution 1 was drained and the gel was agitated in solution 2 (denaturation) and solution 3 (neutralization) for 30 minutes each on an orbital shaker at room temperature.

The blotting apparatus was prepared by placing two layers of 3 MM Chromatography Papers (Whatman, Maidstone, UK) wet with 20× SSC buffer (Appendix 2) on top of a glass plate with both edges of the papers soaked in the 20× SSC buffer. A sheet of gladwrap was placed over the filter papers, and cut using a scalpel to leave a “window” approximately 1-2 mm smaller than the size of the gel. This prevents the flow of SSC by-passing the gel when the blot is assembled. A small amount of 20× SSC was poured on top over the top filter paper, and the gel was placed upside down on the paper. The gel was covered with 20× SSC, and a Hybond-N+ membrane (GE Healthcare Ltd) was placed on the gel. Three more layers of wet Chromatography Papers (Whatman, Maidstone, UK) were placed over the membrane by making sure there was no air bubble left between each of the layers. Finally, two stacks of paper towels and 500 g weight was placed on top of the apparatus and it was allowed to stand at room temperature overnight. The following day, the membrane was washed in 2× SSC buffer and air-dried. Before hybridization, the membrane was cross-linked with a UV crosslinker (Cex-800 UV crosslinker, Ultra-Lum Inc).
2.10.3. Hybridization of DIG labeled probe

The membrane from Section 2.10.2 was placed in a hybridization tube with DNA side facing inwards, and soaked in DIG Easy Hyb solution (10 ml/100 cm² of membrane; Roche, Penzberg, Germany) by rotating in a Bachofir hybridization oven at 45°C for 2 hours. This prehybridization step reduces background noise by blocking non-specific nucleic acid binding sites of the membrane. The probe was denatured in boiling water for 10 minutes, and then rapidly cooled on ice. The denatured probe, in a concentration determined by the method mentioned in Section 2.10.1, was added to the hybridization buffer, and rotated overnight at 45°C. Next day, the membrane was washed twice in washing solution 1 (Appendix 2) for 5 minutes in room temperature. Finally, the membrane was washed twice in washing solution 2 (Appendix 2) for 15 minutes at 68°C for a high stringency wash to allow only highly homologous sequences to remain bound by the probe. The hybridization buffer containing probes was stored at -20°C for later use if needed.

2.10.4. Signal detection

After hybridization and stringency washes, the signal was detected according to the instructions in DIG Application Kit Manual for Filter Hybridization (Roche). The membrane was washed in Buffer I (Appendix 2) for 2 minutes, and then incubated in 30 mL Buffer II (Appendix 2) for 30 minutes with shaking. Anti-Digoxigenin-AP antibody (Roche, Penzberg, Germany) was centrifuged at 16,200g for 5 minutes in order to avoid background that can be caused by small antibody aggregates, and added to buffer II in a ratio of 1:10,000, then the membrane was incubated for another 30 minutes with shaking. The membrane was washed for 15 minutes twice in 150 mL Buffer I with shaking, followed by equilibration for 5 minutes in Buffer III (Appendix 2). The membrane was placed in an A4 copy-safe pocket with DNA side facing up, and approximately 1 mL of CSPD® ready-to-use chemiluminescence substrate (Sigma-Aldrich, Missouri, USA) was applied to the membrane by gently spreading across the membrane without leaving bubbles. After incubating at 37°C for 10 minutes to activate the enzyme, the membrane was
exposed to X-ray film (Fuji film) for up to 3 hours. Finally, the film was developed using a 100 plus Automatic x-ray Processor (All Pro imaging).

2.11. Secondary metabolite extraction, detection, quantification

2.11.1. Extraction of metabolites

*D. septosporum* mycelia, grown as in Section 2.2.2, were separated from broth using an autoclaved funnel and nappy liners. For the determination of different media conditions (Section 3.1.3), metabolites from mycelia or 5 mL of broth were extracted by agitation at 180 rpm in 5 mL acetone or ethyl acetate (EtAc), respectively, for 3 days at 30°C in the dark.

For the solvent optimisation for Nps3 secondary metabolite detection (Section 3.2.3.6), metabolites were extracted from culture filtrate using one of three methods: 1) Ethyl acetate (EtAc), 2) chloroform then dichloromethane, or 3) dichloromethane then 1-butanol. Metabolites were extracted from mycelia using: 1) acetone, 2) dimethylformamide (DMF), then dimethyl sulfoxide (DMSO), then acetonitrile, or 3) Isopropanol, then ethanol (EtOH), then methanol (MeOH). Different solvents were used sequentially for methods 2 and 3 in each case in order to visualise the broadest spectrum of metabolites on each TLC plate compared to the use of a single solvent. Due to the limited availability of water-insoluble solvents with high polarity, fewer solvents could be used for the analysis of culture filtrates compared to mycelia.

Following extraction of metabolites from mycelia and culture filtrates with agitation in the organic solvents, the mixtures were centrifuged at 5320 g for 5 minutes and the organic solvents were separated from the mycelia or broth. From the metabolite-containing solvents, an amount between 200 μL to several mL as required was transferred and vacuum dried using a rotary evaporator (Watson Victor, Australia) or transferred to a microcentrifuge tube and left to dry in a fume hood overnight. The dried metabolites were resuspended in 20 μL acetone or ethyl acetate.
2.11.2. Thin Layer Chromatography (TLC)

For the analysis of *D. septosporum* secondary metabolites, TLC analysis was performed by using TLC Silica Gel 60 plates (Merck, NJ, USA). To determine the effects of different growth media on SM profiles, 200 μL of the metabolites prepared as in Section 2.11.1 were used. For comparison of SMs extracted from the ΔDsNps3 and wild type *D. septosporum* strains, the amounts of SM extracts loaded onto the TLC were normalized based on dothistromin levels in the samples estimated by measuring intensity of TLC bands using Gel-Doc™XR documentation system (Bio-Rad) after one round of preliminary TLC. The TLC was performed in a glass chamber with 1 cm depth of 60:40:0.1 toluene:acetone:formic acid solvent until the mobile phase had moved 18 cm. The TLC plate was then removed from the glass chamber and air dried in the dark in a fume hood. Finally, the TLC plate was visualized on a UV transilluminator (Alpha Innotech) and photographed using a Nikon D7000 camera with an AF Micro Nikkor 60mm f2.8 lens at the Manawatu Microscopy and Imaging Centre (MMIC).

2.12. Phenotypic characterization of WT and ΔDsNps3 *D. septosporum*

2.12.1. Pathogenicity assay

The pine seedling inoculation method was adapted from Kabir et al. (2013). For the pathogenicity assay, wild type, ΔDsNps3 KO1 and ΔDsNps3 KO2 *D. septosporum* spores were prepared as in Section 2.2.2. Spore numbers and morphology assessments were performed using a Zeiss 47 06 00-9901 microscope and their concentrations determined using a haemocytometer.

The spores were diluted to 2 x 10^6 spores/mL with sterile ddH₂O. Four replicate pine seedlings that are susceptible to *D. septosporum* with varying resistance levels, named P1-P4 (Section 2.1.3) were inoculated with approximately 30 mL spore suspension for each strain, using a McGregor's Pressure Sprayer. However, only two seedlings (C and D) could be inoculated with ΔDsNps3 KO1 due to insufficient spores. After the needles were air-dried, the individual seedlings were
covered with plastic autoclave bags and placed in 90 L plastic containers (39.5 cm x 69 cm x 45 cm) containing approximately 20 L distilled water in which two misters were immersed to provide high moisture throughout the chamber (Kabir et al., 2013). The plastic containers were also covered with a frame covered with plastic wrap, with a small opening on the top to adjust moisture. The individual covers on the seedlings were removed after one week and the positions of the plants within each plastic box were rotated clockwise every 3-4 days to facilitate even light and moisture distribution. The experiment was conducted in natural light conditions.

The needles were harvested from the seedlings at 18 weeks post-inoculation (wpi). The lesions were then identified and counted using a Leica MZ10F binocular microscope (Leica Microsystems, Germany) fitted with a Leica DFC 450C digital camera using Leica Application Suite (LAS) v3.8, then measured. The number of lesions used in this experiment is shown in Table 3.6.

2.12.1.1. Surface hyphal network determination

For the determination of the extent of surface hyphal network on wild type and ΔDsNps3 *D. septosporum* infected pine needles, two random needles from the top, middle and bottom of each tree each were picked at 15 days post inoculation (dpi) and soaked in absolute ethanol until use. The needle sections were cut longitudinally and soaked in solution A (1:3 acetic acid: ethanol) overnight, followed by solution B (1:5:3 acetic acid: ethanol: glycerol) for 3 hours with shaking at 180 rpm. Using a method adapted from Mehrabi et al., (2006), the cleared needle sections were stained in 0.01% trypan blue in lactophenol (40% glycerol, 20% lactic acid, 20% phenol in ddH₂O) in a microcentrifuge tube for 15 minutes in boiling water, and destained in saturated aqueous solution of chloral hydrate (VWR International, PA, USA) (5:2, wt/vol) overnight. The stained needle sections were then stored in 87% glycerol at room temperature. Trypan blue stained samples were visualised using a fluorescence microscope. First, needle sections were mounted on microscope slides in 87% glycerol. Then, analysis was performed by using an Olympus BX51 fluorescence microscope (Olympus
Corporation, Japan) fitted with Qimaging micropublisher 5.0 RTV. Fluorescence of trypan blue stained hyphae was done using a triple filter with excitation ranges at 380-400 nm, 475-490 nm, 540-565 nm and with emission ranges at 455-470 nm, 510-530 nm, 590-620 nm, respectively. The images were taken using Q-Capture Pro7 and processed using ImageJ (Schneider et al., 2012).

2.12.1.2. Biomass estimation

Standard curve analysis and estimation of fungal to plant biomass ratio were done using qPCR. qPCR reaction mixes were prepared using SensiFAST™ SYBR No-ROX Kit (Bioline Reagents, UK) according to the manufacturer's instructions and reactions were carried out in LightCycler® 480 (Roche). *D. septosporum* AflJ (target gene) and *P. radiata* cinnamyl alcohol dehydrogenase (CAD) (reference gene) primer pairs (Appendix 3) were used to determine fungal and plant biomasses, respectively. First, serial dilutions of *D. septosporum* (10, 2, 0.4, 0.08, 0.016, 0.0032, 0.00064 ng/μL) and *P. radiata* (50, 10, 2, 0.4, 0.08, 0.016, 0.0032 ng/μL) pure gDNAs in ddH₂O were made in order to prepare standard curves. In three technical replicates, 2 μL of gDNA suspensions were mixed with 8 μL reaction mix and 50 cycles of PCR reactions were carried out (5 sec at 95 °C, 10 sec at 60 °C and 20 sec at 72 °C) with a detection temperature of 72 °C. The cycle number in which the fluorescence could be distinguished from background for the first time is the cycle threshold (Ct) value. A standard curve was prepared by plotting the logarithm of the 5-fold serial dilutions of fungal and plant gDNAs against the Ct values.

For the estimation of fungal to plant biomass ratio, Dothistroma disease lesions harvested from needles at 18 wpi (Section 2.12.1) were cut and freeze-dried. Then, gDNAs were extracted from these lesions using a Genomic DNA Mini Kit (Plant) according to the manufacturer’s specifications. The relative amounts of fungal to plant biomass in each sample were calculated from the results of qPCR amplification of fungal (AflJ) and plant (CAD) genes in that sample using RealQuant® version 1.1.1 (Roche). For each of the technical replicates, ΔDsNps3 to wild type estimated biomass ratios were calculated separately and their averages and standard deviations were used in the results.
2.12.2. Sporulation rate analysis

*D. septosporum* wild type and ΔDsNps3 spores in DSM were harvested as in Section 2.2.2 and diluted to 10^6 cells/mL. From these, 100 μL were inoculated into each of three replicate PMMG plates. After 1 week of growth in the dark, 2 mL sterile ddH₂O was added to each plate, spread with a sterile glass spreader and incubated for 10 minutes to enable the release of the spores. Afterwards, the spore-containing water was transferred to a microcentrifuge tube using a sterile glass spreader and 1 mL sterile pipette. Finally, the spore concentrations were estimated using a haemocytometer by counting two technical replicates for each spore suspension.

2.12.3. Radial growth rate analysis

From *D. septosporum* wild type, ΔDsNps3 KO1 and KO2 7 day old cultures grown in DM, two 5 mm diameter agar plugs with mycelia were extracted using a sterile cork borer and transferred to each of four replicate DM plates for each strain. Horizontal and vertical colony diameters were measured every 3 days for 30 days.

2.13. Statistical analyses

Sporulation, radial growth rates, and estimated biomass ratios of wild type and ΔDsNps3 were compared with student’s t-test using Microsoft Excel to test the hypothesis that there was no difference between wild type and ΔDsNps3 strains at p≤0.05. For the statistical analysis of the effects of different media conditions, multiple t-tests at p≤0.05 were corrected with the false discovery rate (FDR) approach using a two-stage linear step-up procedure to determine which p values are small enough to investigate further (Benjamini and Hochberg, 1995; Benjamini et al., 2006). For all the statistical analyses, p≤0.05 was assumed significant.
2.14. In silico analyses

2.14.1. Domain analysis

Domain predictions were made using PKS/NRPS Analysis Web-site (http://nrps.igs.umaryland.edu) and confirmed by using Structure Based Sequence Analysis of Polyketide Synthases (SBSPKS) (http://202.54.249.142/~pksdb/sbspks/search_pks_nrps.html), and NCBI domain finder (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).

2.14.2. Secondary metabolite product prediction

SM product predictions were made based on Bradshaw et al. (2006) and de Wit et al. (2012) (Table 3.1) or Natural Product Domain Seeker (NaPDoS) analysis (http://napdos.ucsd.edu/run_analysis.html) which predicts based on comparing KS or C domains of PKS and NRPSs with reference genes from well-characterised chemical pathways.

2.14.3. Phylogenetic analyses

For the phylogenetic analyses of DsPks1, DsPks2, and DsNps3, best FASTA hits (Pearson and Lipman, 1988) were obtained from each of the species belonging to all fungi from the Dothideomycetes, Eurotiomycetes, and Sordariomycetes classes within the JGI MycoCosm database (Grigoriev et al., 2012). Sequences that were also confirmed by best reciprocal FASTA hit back to the *D. septosporum* NZE10 protein models were selected with the help of Dr Pierre-Yves Dupont (Massey University). Initial phylogenetic trees of DsPks2 were built by aligning DsPks2 and its best reciprocal BlastP hits of Dothideomycetes and *Aspergillus* spp. in the JGI database with ClustalW algorithm (Thompson et al., 1994) using full protein sequences, KS domains or KS+AT domains, and by building phylogenetic trees using a neighbour joining method (Saitou and Nei, 1987) in Geneious v8.0.3 (Kearse et al., 2012). The support of the branches was computed using the
bootstrap method with 10000 replicates. Only branches showing a minimum 50% bootstrap support were kept in the final tree.

In order to obtain more reliable phylogenies, the final phylogenetic tree of DsPks2 as well as phylogenetic trees of DsPks1 and DsNps3 were built by aligning the amino acid sequences using the Fast Fourier Transform (MAFFT) multiple alignment program with the E-INS-I set of parameters which is the recommended method for aligning sequences that vary a lot (Katoh et al., 2005). Phylogenetic trees were built using the maximum likelihood method as in PhyML software with default parameters (Guindon et al., 2010). The phylogenetic trees were visualized and coloured with ETE3 toolkit (Huerta-Cepas et al., 2016).

Color-coding of the final phylogenetic trees was done by Dr Pierre-Yves Dupont with scripts available on request. Details of the genome sequences used in all phylogenetic analyses are presented in Appendix 23.

2.14.4. Comparison of SM core genes from 19 strains of *D. septosporum*

Nucleotide sequences of the SM core genes were aligned using MUSCLE (Edgar, 2004) in Geneious v8.0.3 (Kearse et al., 2012) with a maximum of 8 iterations. Positive and negative selections on the SM core genes were determined by Andre Sim using dN/dS ratios calculated using CodeML which is part of Phylogenetic Analysis by Maximum Likelihood (PAML) (Yang, 2007). Parameters for PAML were runmode = -2, seqtype=1, CodonFreq=0, model = 0, NSsites=0, icode =0, fix_kappa=1, kappa=1, fix_omega = 0, omega = 0.5. Codons under statistically significant positive selection were detected using a sitewise likelihood-ratio (SLR) method by Dr Pierre-Yves Dupont (Massingham and Goldman, 2005). First, a phylogenetic tree of the *D. septosporum* strains was built based on a concatenation of all SNPs present in the 19 *D. septosporum* genomes. The concatenation process produced a multiple sequence alignment. The phylogeny was then rebuilt using PhyML with its default parameters for a nucleotide tree. The tree was then used as a base for the computation of the selection profiles with the SLR method.
2.14.5. Gene cluster analyses

For the gene cluster analysis, 20 genes upstream and downstream of each SM core gene were selected and their functions were predicted by a combination of GO terms analysis of *D. septosporum* NZE10 gene sequences from Bradshaw et al. (2016) and manual InterProScan analysis (Jones et al., 2014). Genes predicted to be part of the putative gene clusters were determined by submitting 20 protein sequences upstream and downstream of the SM core gene together onto antibiotics & Secondary Metabolite Analysis SHell (antiSMASH) (Weber et al., 2015) and by searching the literature for genes with similar predicted functions in other SM gene clusters.
3. Results and Discussion

3.1. Secondary metabolites in *Dothistroma septosporum*

3.1.1. Overview of *Dothistroma septosporum* secondary metabolite genes

*D. septosporum* SM genes were identified as a result of two separate studies (Bradshaw et al., 2006; de Wit et al., 2012). The PKS core gene for the biosynthesis of dothistromin, a toxin similar to an aflatoxin precursor, was identified by Bradshaw et al. (2006) and named *DsPksA*, indicating similarity to the orthologous aflatoxin PKS gene. The other *D. septosporum* SM core genes were identified in the genome sequence by de Wit et al. (2012), and these genes were named with numbers such as *DsPks1* and *DsNps2* based on the convention for Dothideomycete genes. There are five PKS genes, three NPS, two HPS and one DMA gene present in *D. septosporum* (Table 3.1).

An overview of *Dothistroma septosporum* SM genes was prepared to summarize the key information about each gene including analyses of the predicted JGI gene models to determine their accuracy, and predictions of gene product types in order to determine the possible functions of the corresponding SMs (Table 3.1).
Table 3.1. Complete list of predicted *Dothistroma septosporum* NZE10 secondary metabolite core genes.

<table>
<thead>
<tr>
<th>Gene</th>
<th>JGI protein ID</th>
<th>Scaffold</th>
<th>Position in scaffold/direction</th>
<th>Number of introns</th>
<th>aa number</th>
<th>Domains</th>
<th>PKS/NPS type</th>
<th>Predicted pathway product type</th>
<th>JGI gene model correct</th>
<th>C. fulvum orthologs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pks1</td>
<td>48345</td>
<td>12</td>
<td>597907-605210 (-)</td>
<td>2</td>
<td>2399</td>
<td>SAT-KS-AT-ACP-ACP-ACP-TE</td>
<td>Type 1/NR-PKS</td>
<td>Dothistromin</td>
<td>Yes</td>
<td>PksA: 194256</td>
</tr>
<tr>
<td></td>
<td>47338</td>
<td>10</td>
<td>127527-134096 (+)</td>
<td>0</td>
<td>2189</td>
<td>SAT-KS-AT-ACP-ACP-TE</td>
<td>Type 1/NR-PKS</td>
<td>Melanin/elsinochrome</td>
<td>Yes</td>
<td>Pks1: 191425</td>
</tr>
<tr>
<td>Pks2</td>
<td>73814</td>
<td>8</td>
<td>910134-918249 (+)</td>
<td>10</td>
<td>2507</td>
<td>KS-AT-DH-ER-KR-ACP</td>
<td>Type 1/HR-PKS</td>
<td>Fumonisin</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Pks3</td>
<td>90367</td>
<td>8</td>
<td>974279-982774 (-)</td>
<td>10</td>
<td>2126</td>
<td>KS-AT-DH-ER-KR</td>
<td>Type 1/HR-PKS</td>
<td>Epothilone</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Pks4</td>
<td>72190</td>
<td>5</td>
<td>2140446-2146542 (-)</td>
<td>1</td>
<td>421</td>
<td>Pseudogenised</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nps1</td>
<td>52251</td>
<td>3</td>
<td>2664965-2682846 (+)</td>
<td>2</td>
<td>5088</td>
<td>C-A-T-C-A-T-C-A-T-C-A-T-C</td>
<td>Multimodular</td>
<td>Cyclosporin</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Nps3</td>
<td>71189</td>
<td>4</td>
<td>1327192-1335024 (+)</td>
<td>0</td>
<td>2610</td>
<td>A-T-C-A-NM-T</td>
<td>Multimodular</td>
<td>Cyclosporin</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Hps1</td>
<td>180045</td>
<td>11</td>
<td>965798-977722 (+)</td>
<td>0</td>
<td>3974</td>
<td>KS-AT-DH-CM-KR-C-A-T</td>
<td>Compactin/Bacitracin</td>
<td>Compactin/Tyrosidine</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Hps2</td>
<td>157678</td>
<td>9</td>
<td>914615-927111 (-)</td>
<td>7</td>
<td>4029</td>
<td>KS-AT-DH-CM-KR-C-A-T</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dma1</td>
<td>28625</td>
<td>11</td>
<td>963990-965313 (-)</td>
<td>1</td>
<td>420</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


b Position [nt] of predicted gene in scaffold according to JGI.


d Types of PKS and NRPS based on predicted domain structures. NR-PKS: Non-reducing PKS, HR-PKS: Highly reducing PKS.

e Predicted pathway product types based on NaPDoS analysis or previous studies (Section 2.14.2). *DsDma1* is adjacent to and divergently transcribed from *DsHps1*, therefore might be part of the same SM cluster (Sections 3.1.2 and 3.4.1.3).

f Gene model is correct according to this study.

g *C. fulvum* orthologs of *D. septosporum* SM core genes based on de Wit et al. (2012); numbers shown are JGI protein IDs.

* Indicates a pseudogene.
The results presented in Table 3.1 show that proteins encoded by all \textit{D. septosporum} \textit{Pks} genes, like most fungal PKSs, are type 1, among which \textit{DsPksA} and \textit{DsPks1} are non-reducing while \textit{DsPks2} and \textit{DsPks3} are highly reducing polyketide synthases. All proteins encoded by \textit{D. septosporum} \textit{Nps} genes are multimodular, containing repetitive A, T, and C sites, but it is not yet known whether each domain is used once or multiple times. \textit{D. septosporum} has one gene predicted to encode an indole alkaloid dimethylallyl tryptophan synthase (\textit{DsDma1}), and this gene is adjacent to \textit{DsHps1}. Therefore, analyses on \textit{DsDma1} were done together with \textit{DsHps1} in Section 3.4.1.3.

Among \textit{D. septosporum} SM core genes, \textit{DsPksA} is the only characterised and extensively studied one, responsible for the biosynthesis of dothistromin (Bradshaw et al., 2006; Zhang et al., 2007; Schwelm and Bradshaw, 2010; Chettri et al., 2013). In addition to \textit{DsPksA}, \textit{DsPks1} and \textit{DsNps2} also have \textit{C. fulvum} orthologs, and these are predicted to be involved in melanin and siderophore biosynthesis respectively (de Wit et al., 2012). The predictions for other \textit{D. septosporum} SM core genes are less certain. Their predicted pathway product types were based on KS domain amino acid sequences for PKS, and C domain amino acid sequences for NPS, by using Natural Product Domain Seeker (NaPDoS) (Ziemert et al., 2012), which estimates what type of SM might be associated with these domains (Section 2.14.2). For example, NaPDoS analysis of \textit{DsPksA} predicts the pathway product as aflatoxin, which is a close prediction as dothistromin is structurally similar to an aflatoxin precursor, versicolorin B (Bradshaw et al., 2006). The actual SM types may be different from the predictions because of two main limitations. First, NaPDoS prediction is based only on one domain in the core gene and does not take into account the genes within the possible gene cluster (Section 1.4.5). The other limitation is that the prediction is based on a database of known SMs, and therefore limited to the SMs within the database. Therefore, true identification of the \textit{D. septosporum} SMs requires extraction of metabolites and chemical analysis. Predictions for the pathway products of the hybrid PKS-NPS genes \textit{DsHps1} and \textit{DsHps2} have two different SM types as the analysis tool can only predict PKS or NRPS using KS or C domains respectively, and Hps have both of these domains.
3.1.2. **Expression in planta**

To help determine if any of the *D. septosporum* SMs might have a role in the disease process, expression *in planta* was assessed. During this project, transcriptome analysis at early (hyphal surface network and early penetration), mid (early lesion) and late (sporulating lesion) stages of infection on *Pinus radiata* needles was carried out (Bradshaw et al., 2016) (Figure 1.3a). As a result of transcriptome analysis, more than three million reads were obtained, the majority of which belonged to the plant host. From early to late stage of infection, the ratio of fungal reads over plant reads increased from 0.1% to 17.1% (Table 3.2) (Bradshaw et al., 2016). This showed that there was a rapid increase in the percentage of fungal reads between mid to late stages, which was consistent with a high rate of biomass increase during the course of infection at this stage (Kabir et al., 2015).

**Table 3.2.** Percentage of fungal reads *in planta* during early, mid, and late stages of infection.

<table>
<thead>
<tr>
<th>Stage of infection</th>
<th>Fungal reads % of total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early</td>
<td>0.1</td>
</tr>
<tr>
<td>Mid</td>
<td>0.5</td>
</tr>
<tr>
<td>Late</td>
<td>17.1</td>
</tr>
</tbody>
</table>

Data obtained from Bradshaw et al. (2016).

Genome-wide expression analysis of *D. septosporum* during the infection process has given insight into many aspects of the fungus and the plant-fungus interactions, but the main focus in this project was the expression of secondary metabolism genes during infection (Figure 3.1).
Figure 3.1. Expression of core SM genes in *D. septosporum* during early, mid, and late stages of infection. RPMK: Fungal reads per million per kilobase. Significant differences (p<0.05) between early-mid and mid-late stages are marked with “*” and “+”, respectively. The differences between early and late stages were all also significant (p<0.05) except for *DsHps1* and *DsHps2*. The gene expression levels could not be validated by qRT-PCR because of the low fungal read proportions for early- and mid-stages of infection.

The analysis of expression data requires setting a threshold for identifying background expression. Since *DsPksA* is known to be involved in production of dothistromin with the expression level represented by approximately 100 RPMK, expression levels of the other SM core genes less than 10% of that of *DsPksA* (i.e. 10 RPMK) were considered to have background-level expression that is most likely not biologically relevant. In cases where higher levels of gene expression are seen, the stage where a SM core gene is most highly expressed can be crucial in determining the possible role of its gene product. In the hemibiotrophic plant pathogens *Colletotrichum orbiculare* and *C. higginsianum*, most of the key SM genes were highly expressed at pre-penetration and biotrophic stages of infection (O’Connell et al., 2012; Gan et al., 2013). Since *DsPks4* is a non-functional gene (Table 3.1), it was not included in the expression analysis. Among the other 10 core SM genes of *D. septosporum*, only *DsNps3* had significantly higher expression level at early stage compared to both other stages, while *DsHps2* expression level at early stage was significantly higher than that of mid stage, but not than that of late stage. These results suggest a possible role for *DsNps3* and *DsHps2* in the biotrophic phase of infection. In addition, *DsPksA* and *DsDma1* were expressed at a significantly higher level at mid stage and *DsPks1* and *DsPks2* had significantly...
higher-level expressions in late stage compared to the other stages. Therefore, it is possible *DsPksA, DsPks1, DsPks2* and *DsDma1* might have a role in the necrotrophic phase of infection. *DsPks3, DsNps1, DsNps2* and *DsHps1* were expressed at a very low level and below the threshold RPMK value; therefore they might not have a role during plant infection. Among *D. septosporum* SM genes, the genes with highest expression at early and late stages, *DsPks1, DsPks2* and *DsNps3*, were mainly focused in this project, and *DsNps3* was selected for functional analysis.

### 3.1.3. Effects of media conditions on secondary metabolism

Fungal secondary metabolite production is usually a complex and energy-exhaustive process (Calvo et al., 2002). Therefore, the SM production is generally tightly controlled to avoid resource consumption where it is not beneficial (Gacek and Strauss, 2012). Expression of *D. septosporum* SM core genes were analysed not only *in planta*, but also in culture (Bradshaw et al., 2016). Apart from *DsPksA, DsPks1, DsPks2*, and *DsNps3*, expression of *D. septosporum* SM genes were lower than the threshold value, 10 RPMK, under rich media conditions (DM).

Production of SMs in fungi can be altered by many biotic and abiotic factors (Section 1.5). In this project, carbon and iron availability as well as nitrogen source were selected as modification targets to determine if any changes occurred in the SM profiles of *D. septosporum*. Carbon availability and nitrogen source were selected because changes in these conditions are known to affect SM production in other fungi (Barboráková et al., 2012, Arai et al., 2012). Iron availability was studied in order to examine its effects on biosynthesis of a predicted siderophore, a type of SM mainly specialized for iron transport and storage (Neilands, 1995). In addition to using a wild type strain, a global SM regulator mutant Δ*LaeA* was used in order to observe any changes in some SMs that are not normally produced at high levels in the wild type strain (Section 1.5).

Figure 3.2 shows TLC analysis of *D. septosporum* metabolites extracted from culture filtrate and mycelia. Since SMs may have roles inside or outside the cell, they may be secreted or kept within the cell (Martín et al., 2005). Therefore, in
order to observe a broader spectrum of *D. septosporum* SMs, both culture filtrate and mycelia were analysed using TLC. It was shown that one band (M3) from the TLC analysis of mycelia was not visualised in culture filtrate, and one band from culture filtrate (F6) was observed at a lower level in the mycelia.
Figure 3.2. TLC analysis of *D. septosporum* wild type and Δ*LaeA* grown in different conditions. SMs from a) Culture filtrate and b) Mycelia of Δ*LaeA* (LaeA) and wild type (wt) strains. C: PMMG medium with 1% glucose, Fe: PMMG medium with no iron, NH₄ and NO₃: PMMG media which contained these compounds used as sole nitrogen sources, (+):PMMG medium with 3% glucose, 10 μM FeSO₄·7H₂O, and NH₄NO₃ as nitrogen source. Bands of interest are marked F1-F6 for culture filtrates and M1-M5 for mycelia. Dothistromin bands are marked with F3 and M4.
In Figure 3.2, the known metabolite dothistromin was evident in bands F3 and M4. The overall amount of metabolites produced was increased in the ΔLaeA strain compared to the wild type in all lanes. Since the expression of D. septosporum SM genes in the ΔLaeA strain was previously reported (Chettri and Bradshaw, 2016) and illustrated in Figure 3.3a, it was possible to compare the TLC bands with expression profiles to determine if any TLC bands might correspond to any specific SMs. Expression of DsPksA, DsHps1, and DsHps2 were significantly increased, and DsPks2, DsNps1 and DsNps2 were significantly decreased in ΔLaeA compared to wild type D. septosporum (Figure 3.3a). Based on the selected TLC bands and expression of SM genes in ΔLaeA (Chettri and Bradshaw, 2016), which bands might belong to which SMs were estimated for the other SMs (Figure 3.3).

Quantification of the TLC data was performed using JustTLC Version 4.0.3, Sweday (Lund, Sweden) (Section 2.11.2). In this analysis both absolute intensity of the bands (non-normalized) and relative intensity of each band over total intensity per sample (normalized) were used. Normalisation was done because of the variations in efficiency of metabolite extraction and the amounts loaded onto the TLC plates. Non-normalized data were also analysed in order to avoid bias when levels of specific SMs changed so much between samples (such as the increase in band M1 in the ΔLaeA mutant) that it affected the total amount of metabolite per sample, thus affecting the normalization. For example, expression of the dothistromin core gene DsPksA was significantly increased in ΔLaeA compared to the wild type. However, a significant increase of band intensity (F3) was only observed in metabolites extracted from culture filtrate and analysed using non-normalized data (Figure 3.3b). Therefore, significant differences from either non-normalized or normalized data were accepted unless the results were contradictory with each other for these two analyses e.g. significant increase in non-normalized and decrease in normalized analysis.
Figure 3.3. Comparison of SM gene expression and metabolite band intensity in TLC analysis between *D. septosporum* wild type and Δ*LaeA* strains. Wt: Wild type, KO: Δ*LaeA*. a) SM gene expression in Δ*LaeA*; data from Chettri and Bradshaw (2016) at 9 dpi. b) Non-normalized and normalized TLC bands from culture filtrate, c) Non-normalized and normalized TLC bands from mycelia. All cultures were grown in PMMG. TLC bands belonging to mycelial metabolite extracts are named M1-M5, and the ones in culture filtrate are named F1-F5. The dothistromin band is indicated. Asterisks indicate significant differences (p<0.05) between wt and Δ*LaeA*. Data are means and standard deviations of two biological replicates from 15 dpi cultures.
According to the results presented in Figure 3.3, expression of *DsHps1* increased most dramatically in the Δ*LaeA* strain compared to the wt. Since the metabolite with highest retention factor (Rf) value in both culture filtrate and mycelia (F1 and M1) appears absent in wild type but produced at a very high level in Δ*LaeA* it was initially thought this could be the product of Hps1. However, recent TLC analyses made by Dr Pranav Chettri showed that the F1/M1 metabolite was still produced in a *DsHps1* and *DsLaeA* double-knockout mutant, while it was absent in a *DsPksA* and *DsLaeA* double-knockout mutant. Therefore, the metabolite presented as F1 and M1 is probably an intermediate of the dothistromin biosynthesis pathway catalysed by DsPksA, which was also upregulated in Δ*DsLaeA* (Figure 3.3a).

The metabolite present in culture filtrate F6 was also produced at higher levels in Δ*LaeA* compared to wild type on non-normalized data (Figure 3.3b). Since the expression of *DsHps1* and *DsHps2* were significantly increased in Δ*LaeA* compared to wild type (Figure 3.3a), it is possible that this metabolite could be the product of Hps1 or Hps2. However, there were limitations with this analysis. First, the difference between wild type and Δ*LaeA* was not significant for F6 and there was no visual difference in normalized data. In addition, it is possible the metabolite extraction method used was not suitable for the extraction of Hps1 and Hps2 SMs, and that may be the reason for no better candidate being found for these SMs in the TLC analysis. On the other hand, expression of *DsPks2* and *DsNps2* were significantly decreased in Δ*LaeA* compared to wild type. However, no band was observed in the TLC analysis that was at a lower intensity for Δ*LaeA* than wild type. Using solvents with varying polarity to extract metabolites may reveal other candidates.

In any case, these were only estimations based on the production of SMs and expression of the core genes, and chemical characterization is required to confirm candidate metabolites. It should also be noted that the expression data presented belong to 9 dpi cultures while the metabolites belong to 15 dpi cultures. Therefore, the expression of the SM genes at 15 dpi might be different than at 9 dpi in Δ*LaeA*, causing a limitation for comparison.
Using the same TLC data shown in Figure 3.2, the effect of different media on the production of metabolites was analysed and is shown in Figure 3.4.
Figure 3.4. Secondary metabolite production under different media conditions. C: 1% glucose, Fe: Iron-deficient medium, NH₄: (NH₄)₂SO₄ as sole nitrogen source, NO₃: KNO₃ as sole nitrogen source, PMMG: Control medium with 3% glucose, 10 μM FeSO₄·7H₂O, and NH₄NO₃ as nitrogen source. Secondary metabolites extracted from culture filtrate or mycelia were indicated. Significant differences (p<0.05) between PMMG and different media conditions were marked with "*". The bands F1 and M1 were missing from the wild type samples as mentioned above, and thus not included in the results.

Effects of different media, presented in Figure 3.4, showed that carbohydrate availability had significant effects on the production of some metabolites. Lowering glucose concentration from 3% to 1% in the medium (by comparing wt-C and wt-PMMG) decreased the amount of one metabolite (F2/M2) both in culture filtrate and mycelia of wild type strain (Figure 3.4a-Norm. and b-Both Norm. and Non-norm.). However, this decrease was not observed in the ΔLaeA strain. Instead, the
amount of dothistromin (F3) was significantly decreased in the culture filtrate if the glucose concentration was reduced (Figure 3.4c-Non-norm.). Carbon catabolite repression by the Cys₂His₂ zinc finger global transcription factor CreA is common among filamentous fungi, and thus limiting carbohydrate concentration was expected to increase production of SMs (Reviewed by Yin and Keller, 2011). It is possible there may have been an increase in SM production at a different time point, and extracting the metabolites at 15 dpi might be too late for detecting that increase. For example, dothistromin, unlike most SMs that are mainly produced at late exponential or stationary phase, is produced most at early exponential phase (Schwelm et al., 2008). Therefore, it is possible that the effect of carbohydrate concentration on some SMs might not have been detected.

The effect of two nitrogen sources, NO₃ and NH₄, was tested on metabolite production. There was no significant difference observed in the production of any metabolites if cultures were grown in media containing NO₃ as the sole nitrogen source compared to PMMG. On the contrary, the amounts of one metabolite from culture filtrate (F6) and one from mycelia (M2) were significantly decreased in *D. septosporum* wild type cultures grown in NH₄-containing medium compared to PMMG (Figure 3.4a-Norm. and b-Both). The decrease in the band F6 may be because of the decrease in the secretion or production of the metabolite, or both. In contrast, metabolite M3 (the band that was missing in culture filtrate but clearly visible in mycelia) was significantly increased in the mycelium of wild type cultures grown in media containing NH₄ as sole N source compared to PMMG (Figure 3.4b-Both). In the ΔLaeA strain, two metabolites (F1 and F2) were produced at significantly lower levels in ΔLaeA cultures grown in media containing NH₄ as nitrogen source compared to PMMG (Figure 3.4c-Norm. (F1), Non-norm. (F2)). Levels of the F2/M2 decreased in both wild type and ΔLaeA strains in NH₄-containing medium compared to PMMG, but the decrease in M2 was only significant in wild type (Figure 3.4b-Both) and that in F2 only significant in ΔLaeA (Figure 3.4c-Non-norm.).

Using NH₄ as nitrogen source may either positively or negatively impact the production of fungal SMs. A medium containing NH₄ as sole source of nitrogen was
reported to induce aflatoxin (AF) production in *A. parasiticus* and *A. flavus* (Feng and Leonard, 1998). In addition, the production of veratryl alcohol produced by the fungus *Phanerochaete chrysosporium* significantly decreased upon addition of NH$_4$ to the media (Fenn and Kirk, 1981). On the contrary, sterigmatocystin (ST) production in *A. nidulans* was repressed in medium containing NH$_4$ as sole nitrogen source (Feng and Leonard, 1998). However, Keller et al. (1997) has shown that the pH regulation of AF and ST may override the effects of nitrogen source. Therefore, a change in pH caused by the use of (NH$_4$)$_2$SO$_4$ in the medium containing NH$_4$ as the sole nitrogen source instead of NH$_4$NO$_3$ in PMMG may be the reason of the changes in metabolite production in *D. septosporum*. Even though the molecular mechanism for nitrogen regulation of most fungal SMs has not been reported, nitrogen source and availability is one of the key factors in SM regulation, affecting a broad range of SMs. For example, nitrogen source and availability affected 30 of the 45 predicted SM gene clusters in the fungal plant pathogen *Fusarium fujikuroi*, including broad range of PKS, NRPS, DMA gene clusters (Wiemann et al., 2013). Therefore, it was not possible to determine the types of SMs affected by the nitrogen source in this experiment.

Among the metabolites produced by *D. septosporum* wild type and Δ*LaeA* strains, two were significantly affected by iron availability. Level of one metabolite (F2) was significantly decreased in iron-deficient media compared to PMMG in the wild type strain (Figure 3.4a-Norm.), although not in the Δ*LaeA* strain. A significant decrease was observed in another metabolite (F1) in iron-deficient media compared to PMMG in the Δ*LaeA* strain (Figure 3.4c-Norm.). It is known that iron depletion promotes siderophore production (Johnson, 2008). However, in *C. fulvum* the ortholog of the putative siderophore gene *DsNps2* was barely expressed even in iron depletion conditions (Collemare et al., 2014). Therefore, it is possible that siderophore production *D. septosporum* might not be affected by iron availability.

To sum up, carbon and iron availability as well as use of NH$_4$ as sole nitrogen source, compared to PMMG, resulted in significant changes in the levels of several putative metabolites in the wild type strain. However, these changes did not
provide clues to determine what kinds of metabolites were affected. This was because the changes in the metabolite levels caused by carbon and iron availability were opposite of the expected and nitrogen source may affect the production of a broad range of metabolites. There were several limitations to this study. First of all, normalized and non-normalized data had shown that there was a significant difference in either normalized or non-normalized data, but not supported by the other. Significant differences were accepted in either one of the normalized or non-normalized analyses as long as they were not contradictory. Repeating the experiment with more replicates may determine if these changes are actually significant. Another limitation was the timing of the extraction. Although most fungal SMs are produced at late exponential phase, some, such as dothistromin, might be produced at a higher level at a different phase of growth. Therefore, it is necessary to do these comparisons at different time points to observe effects of different media conditions on a broader range of metabolites. In addition, the method for extraction was another limitation. Since the metabolites extracted may differ based on the polarity of the organic solvent used, it is possible that some of the metabolites may have been extracted insufficiently to observe effects of media conditions with the current method. The final limitation was the pH effects of different media conditions, which may in turn override the effects of other media conditions.

### 3.2. Nonribosomal peptide synthetases

Following on from studies of expression of all of the D. septosporum SM genes, and metabolite production under different media conditions, the genes were studied individually. First, JGI gene models for the D. septosporum SM core genes were confirmed in order to make predictions based on these core genes. Afterwards, evolutionary selection pressures were analysed for the SM core genes to help determine the importance of the gene product and its domains for the fungus. Then, putative SM gene clusters were determined to make predictions about the SM biosynthesis pathway products. Three of the genes (DsPks1, DsPks2, and DsNps3) were of particular interest based on their levels of expression and expression patterns. Additional analyses for the genes of particular interest
included phylogenetic analyses to find out if there were functionally characterised orthologs in other species, and functional characterization of DsNps3, the gene that was up-regulated at an early stage of infection in planta.

### 3.2.1. NPS1

#### 3.2.1.1. Confirmation of gene model for DsNps1

All the analyses and predictions related to the D. septosporum SM core genes are based on the assumption that the JGI gene models are correct. Therefore, it was essential to confirm these gene models for each SM core gene before making any further predictions about these genes and their potential roles. In order to confirm each gene model two methods were used. The amino acid sequences of SM core gene products were aligned with those from their best BlastP hits and known orthologs (if any), and the predicted introns were confirmed by mapping reads from in planta and in culture RNA-seq experiments to the gene model. All of the RNA-seq reads that were mapped to the gene model was uploaded onto JGI by Andre Sim (Massey University) and are available by browsing the genes within the scaffolds from the JGI website.

DsNps1 had no known orthologs, so its amino acid sequence was aligned with those obtained from best BlastP hits that were confirmed using reciprocal BlastP (Figure 3.5). With this approach it was possible to detect ambiguities in gene models caused by factors such as a mis-annotated introns or start/stop codons, which may cause big gaps or insertions in the protein sequence compared to other sequences within the alignment. However, it should be noted that for this comparison to work, the sequences within the alignment should not be too different and preferably orthologs.

Figure 3.5 illustrates the predicted domain structure of DsNps1 (Table 3.1), using an alignment of DsNps1 with the top three matches from a reciprocal BlastP analysis. Alignment of DsNps1 with its closest BlastP hits showed very low consensus, and low amino acid sequence identity with the closest of those being A. wentii at 44.3% (E=0). The sequence identity of DsNps1 with the two other closest
BlastP hits, *T. citrinoviride* and *G. acutata* were 37.1% (E=0) and 35.1% (E=0), respectively. *D. septosporum* belongs to the Dothideomycetes class, however, none of the best BlastP hits belonged to this class. *A. wentii* is within the Eurotiomycetes class, while *T. citrinoviride* and *G. acutata* are Sordariomycetes. The best BlastP hit of DsNps1 from among 102 species of Dothideomycetes was *Botryosphaeria dothidea* (protein ID: 14052) with 31.9% identity. This shows that *DsNps1* is unique to *D. septosporum* amongst fungi whose genomes have been sequenced. Therefore, due to low similarities between DsNps1 and its best BlastP hits, it is not possible to reach a conclusion about the *DsNps1* JGI gene model based on the alignments.

The *DsNps1* predicted gene model has two introns, however, both *in planta* and in culture expression levels of the gene were very low (below 10 RPMK). This meant there weren’t sufficient cDNA reads near the positions of predicted introns to confirm their presence (Appendix 4). However, the second intron is within an A domain, and this region of DsNps1 aligns well with A domains of other proteins within the alignment (Figure 3.5), suggesting the second intron is very likely correct.
Figure 3.5. Domain structure of DsNps1 and its protein alignment with best BlastP hits. DsNps1 indicates the peptide’s domain organization and amino acid positions of some of the domain borders are shown with numbers (T domain borders and borders of domains that are too close to another annotated border are not shown). C: Condensation, A: Adenylation, T: Thiolation domains. Ds: D. septosporum, Aw: Aspergillus wentii (protein ID: 24303); Tc: Trichoderma citrinoviride (protein ID: 1118665), Ga: Glomerella acutata (protein ID: 1396682). The diagram was drawn in scale to original domain sizes. Triangles indicate corresponding intron positions. The dark regions indicate the amino acid alignment with black (highly similar) to white (not similar) sequences with respect to the consensus sequence. Total length is 5888 amino acids. Coloured bar represents consensus identity: red indicates low identity (below 30%); green shades indicate higher levels of identity (up to 100%) with bar height proportional to % identity.
Therefore, neither alignments (Figure 3.5, Appendix 20a) nor cDNA analyses (Appendix 4) provide sufficient evidence to confirm the current gene model. The current gene model should be confirmed by cDNA sequencing of *DsNps1* but, because of its low level of expression, cDNA would have to be obtained using growth conditions, or in a mutant, that results in overexpression of *DsNps1*.

### 3.2.1.2. Comparison of *DsNps1* from 19 strains of *D. septosporum*

Plant pathogens are in a constant evolutionary struggle with host plants in order to gain competitive advantage, and fungal SMs usually play key roles in this (Section 1.4). In this process, genes may be under evolutionary selection. Evidence of evolutionary selection can be inferred from proportions of non-synonymous (amino acid changing, dN) and synonymous (silent, dS) mutations in codon sequences (dN/dS). It is now also possible to determine the selection pressure on the specific codons of a gene by the help of various bioinformatics tools (Aguileta et al., 2009). This can provide a deeper understanding on the evolutionary selection pressure on specific domains of the encoded proteins.

Genomic DNAs of 18 *D. septosporum* strains from across the globe were received from Dr Irene Barnes (University of Pretoria, South Africa) and sequenced in AGRF (Australia) with Illumina sequencing, and assembled to the gene model using Bowtie 2 (version 2.2.6) (Langmead and Salzberg, 2012). Full details of strains are presented in Appendix 5. *DsNps1* evolutionary selection determination was done by aligning the *DsNps1* nucleotide sequences of NZE10 and the 18 additional *D. septosporum* genomes, screening for internal stop codons and calculating overall dN/dS ratio as explained in Section 2.14.4. Figure 3.6 shows the *DsNps1* nucleotide sequence alignment from the 19 *D. septosporum* strains. The dN/dS ratio of *DsNps1* is 0.23, indicating the gene is under negative selection. Therefore, although the *in planta* expression level of *DsNps1* was very low, the Nps1 SM may have an important function at another stage of the *D. septosporum* lifecycle. In addition, this alignment supports that the JGI gene model is correct.
Figure 3.6. *DsNps1* nucleotide alignment of NZE10 with 18 additional strains. Red arrows show the domain positions with, C: condensation, A: Adenylation, Unnamed small red arrows: thiolation domains. Yellow arrows indicate intron positions. The coloured bar represents consensus identity with brown indicating identities above 30% and below 100% and green indicating 100% identity. For each sequence, vertical bars on the top and bottom lines show differences in nucleotide and amino acid sequences compared to NZE10. A bar on the top line alone shows a synonymous mutation and a bar on both lines shows a non-synonymous mutation. Strains used in this analysis are listed in Appendix 5.
In addition to the synonymous and non-synonymous mutations, the alignment presented in Figure 3.6 also shows that DsNps1 genes in COL (Colombia) and GUA (Guatemala) samples have more mutations compared to the other 15 D. septosporum genomes. However, this difference is not specific to DsNps1 and a phylogeny of 19 D. septosporum genomes shows that COL and GUA samples are distant from the other 15 D. septosporum genomes analysed (Figure 3.7).

![Figure 3.7. Phylogenetic analysis D. septosporum whole genome sequences of NZE10 and 18 additional D. septosporum genomes. Numbers indicate approximate likelihood ratio test (aLRT) scores. Phylogeny built by Dr Pierre-Yves Dupont.](image.png)

3.2.1.3. Gene cluster analysis

SM biosynthesis requires combined activity of products of multiple genes that are generally close to each other within a chromosome (Section 1.4.5).
The SM produced by \textit{DsNps1} may not have a key role in virulence based on its low levels of expression \textit{in planta}, but may still function at some stage in the \textit{D. septosporum} life cycle. Gene cluster analysis of \textit{DsNps1} was done by predicting what genes might be associated with \textit{DsNps1} in a gene cluster. This was done by screening 20 genes upstream and downstream of \textit{DsNps1} by GO terms analyses from Bradshaw et al. (2016), manual InterProScan analyses, antiSMASH analyses, and searching the literature for similar genes with predicted functions in other SM gene clusters as outlined in Section 2.14.5. As mentioned above, the \textit{DsNps1} gene has one of the lowest expression levels both in culture and early, mid, and late time stages of infection \textit{in planta} according to transcriptomics analysis (Section 3.1.2). Because of this, it was not possible to include analysis of co-expression between \textit{DsNps1} and the genes nearby on the chromosome to help predict \textit{DsNps1} cluster genes.

Figure 3.8 shows the putative gene cluster for \textit{DsNps1} and Table 3.3 shows the predicted proteins encoded by the genes within the gene cluster. The putative \textit{DsNps1} gene cluster contains 7 genes within a 45 kb region of chromosome 3. In this region, there are 6 additional genes with no predicted functions related to secondary metabolism. The biggest distance between two genes belonging to the predicted gene cluster is 6 kb, containing 1 interfering gene.

Figure 3.8. \textit{DsNps1} and its predicted gene cluster. Numbers inside arrows indicate JGI protein IDs. Pink: Secondary metabolite core gene \textit{DsNps1}. Purple: putative \textit{DsNps1} cluster genes. Small white arrows: genes without predicted SM functions.

<table>
<thead>
<tr>
<th>Protein ID(^a)</th>
<th>Predicted protein(^b)</th>
<th>Best hit organism/JGI ID(^c)</th>
<th>E / identity %</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Ds70757</td>
<td>Major facilitator transporter</td>
<td>\textit{C. fulvum} / 196085</td>
<td>0/85</td>
</tr>
<tr>
<td>Ds169779</td>
<td>O-succinylhomoserine sulfhydrylase</td>
<td>\textit{C. fulvum} / 196087</td>
<td>0/89.7</td>
</tr>
<tr>
<td>*Ds87328</td>
<td>FAD-dependent oxidoreductase</td>
<td>\textit{T. citrinoviride} / 1156674</td>
<td>0/61.9</td>
</tr>
<tr>
<td>*Ds126908</td>
<td>Drug resistance transporter</td>
<td>\textit{C. fulvum} / 188587</td>
<td>7.2e-145/79.4</td>
</tr>
<tr>
<td>*Ds52251/Nps1</td>
<td>Non-ribosomal peptide synthetase</td>
<td>\textit{A. wentii} / 24303</td>
<td>0/44.3</td>
</tr>
<tr>
<td>Ds87332</td>
<td>Carbamoyl phosphate synthase</td>
<td>\textit{W. ornate} / 504389</td>
<td>2e-91/54.5</td>
</tr>
<tr>
<td>*Ds126514</td>
<td>Drug resistance transporter</td>
<td>\textit{Z. cellare} / 64885</td>
<td>0/76.6</td>
</tr>
</tbody>
</table>

\(^a\) \url{http://genome.jgi.doe.gov/Dotse1/Dotse1.home.html}  
\(^b\) Predicted proteins according to antiSMASH and InterProScan analyses.  
\(^c\) All BlastP hits were confirmed with reciprocal blast analyses.
Three of the seven genes within the putative gene cluster are predicted to encode transporter proteins. Transporters can be used to secrete metabolites out of a cell, but are also commonly used to prevent autotoxicity. For example, a transporter gene, *gliA*, in the gliotoxin gene cluster of *A. fumigatus* provides gliotoxin resistance for the fungus itself (Dolan et al., 2015). Although having three transporter proteins seems redundant, there are up to four transporters within single predicted NPS gene clusters of *Trichoderma spp.* (Bansal and Mukherjee, 2016).

Aside from transporters, another gene within the putative gene cluster, *Ds87328*, is predicted to encode a FAD-dependent oxidoreductase. A predicted NPS biosynthetic gene cluster, involved in production of notoamide by a marine-derived *Aspergillus* sp. contains two FAD-dependent monooxygenases (Ding et al., 2010). In addition, *APF9* gene encodes a FAD-dependent monooxygenase in the apicidin F cluster, a gene cluster composed of 11 genes in the rice pathogen *Fusarium fujikuroi*. Deletion of the *APF9* gene completely stopped apicidin F production, but revealed production of its analog apicidin K (Niehaus et al., 2014). *APF9* is thought to catalyse conversion of 2-amino-8-hydroxyoctanoic acid into 2-aminooctanedioic acid, which is then used as a substrate for the NPS core protein of the apicidin F biosynthetic pathway, APF1 (Niehaus et al., 2014).

There are two more genes within the *DsNps1* putative gene cluster; however, no genes with similar functions could be found in other currently-known fungal SM gene clusters. *Ds169779* is predicted to encode an O-succinylhomoserine sulfhydrylase and was included as part of the putative *DsNps1* gene cluster by antiSMASH analysis. This protein catalyzes L-homocysteine production, a step required in methionine biosynthesis (Foglini et al., 1995). Even though no such gene has been reported in SM gene clusters, an O-succinylhomoserine sulfhydrylase gene *metZ* was found to be co-regulated with NPS genes in the pathogenic bacterium *Burkholderia pseudomallei* (Chen et al., 2014). Therefore, it can be hypothesized that a gene predicted to encode an O-succinylhomoserine sulfhydrylase might be part of an NPS gene cluster. The other gene within the predicted *DsNps1* gene cluster, *Ds87332*, is predicted to encode a carbamoyl phosphate synthase, an enzyme required for arginine and pyrimidine biosynthesis.
in prokaryotes and eukaryotes, as well as being involved in urea cycles in vertebrates (Holden et al., 1999; Raushel et al., 1999). Carbamoyl phosphate synthase catalyzes formation of carbamoyl phosphate from ATP, ammonia and bicarbonate (Holden et al., 1999). A gene predicted to encode a carbamoyl phosphate synthase was found in the hybrid PKS-NRPS gene cluster for saxitoxin in the bacterium *Anabaena circinalis* (Neilan et al., 2008). It was also found that carbamoyl phosphate is a substrate required for *in vitro* saxitoxin biosynthesis (Kellmann and Neilan, 2007).

The predicted pathway product type of DsNps1 is cyclosporin (Table 3.1), a compound with a variety of functions including immunosuppression, antifungal and plant virulence activities (Keller, 2015; Viaud et al., 2003). A predicted fungal cyclosporin gene cluster belonging to the fungus *Tolypocladium inflatum* with core gene *simA* (Bushley et al., 2013) was compared with the DsNps1 putative gene cluster. However, genes in the *simA* gene cluster do not have similar functional annotations to the genes in the predicted DsNps1 gene cluster. Instead, the *simA* gene cluster contains genes for cytochrome p450, dehydrogenase, transcription factors and cyclophilin, a protein that binds to cyclosporin. In addition there are also amino acid specific enzymes such as alanine racemase in the *simA* gene cluster. The predicted gene cluster for *simA* is composed of 14 genes, while DsNps1 putative gene cluster is composed of only nine genes. It is possible that other non-clustered genes might be involved in the production of the SM catalysed by DsNps1 similar to dothistromin biosynthesis that is catalysed by enzymes encoded in a fragmented gene cluster as explained in Section 1.4.5.1.

Most commonly found genes in other fungal SM gene clusters include methyltransferase, cytochrome p450 monooxygenase, glutathione S-transferase, acetyltransferase, and short-chain dehydrogenase/reductase (Collemare et al., 2014; Khaldi et al., 2008; Gardiner et al., 2014; Cramer et al., 2006). Gene cluster predictions can only determine which genes might be required in the biosynthesis of a certain SM, however, and actual gene clusters may be very variable. Biosynthesis of peramine, an NRP produced by *Epichloë* spp., is catalysed by a single SM gene, *perA*, without a gene cluster (Berry et al., 2015). The best way to
confirm a gene cluster is to generate gene knockout (KO) mutants for each candidate gene and analyse the loss of metabolite produced.

In conclusion, even though the gene model for \textit{DsNps1} could not be confirmed due to the lack of clear homologs and very low expression levels, \textit{DsNps1} gene alignment across 19 \textit{D. septosporum} strains supported the gene model of \textit{DsNps1}. \textit{DsNps1} has no clear homologs and is under negative selection. The predicted gene cluster for \textit{DsNps1} has 7 genes, however, predicted functions of the genes in the putative gene cluster do not support the prediction that the DsNps1 SM is a cyclosporin.

### 3.2.2. NPS2

#### 3.2.2.1. Confirmation of gene model for \textit{DsNps2}

A misannotated intron or start/stop codon can change the predicted protein and domain structure of a nucleotide sequence dramatically, so confirmation of this gene model was mandatory before doing further analyses. \textit{DsNps2} is one of the three \textit{D. septosporum} SM core genes with a known \textit{C. fulvum} ortholog and is predicted to produce a siderophore (Table 3.1). Gene model confirmation was made by amino acid sequence alignment of \textit{DsNps2} with its \textit{C. fulvum} ortholog and two other best reciprocal BlastP hits, as well as by confirming the predicted introns using in culture and \textit{in planta} expression analysis data.

Figure 3.9 shows the domain organization of \textit{DsNps2} and an alignment of \textit{DsNps2} with its \textit{C. fulvum} ortholog CfNps2 and two other top hits from reciprocal BlastP analysis. Alignments had an overall high consensus, and \textit{DsNps2} amino acid identities (and BlastP probabilities) to \textit{C. fulvum}, \textit{Z. cellare} and \textit{M. fijiensis} were 76.6% (E=0), 57% (E=0), and 55.6% (E=0), respectively. These alignments suggested there are no gaps or insertions present in \textit{DsNps2} compared to the consensus sequence, which supports the hypothesis that the gene model for \textit{DsNps2} is correct.
Figure 3.9. Domain structure of DsNps2 and its protein alignment with best BlastP hits. DsNps2 indicates the peptide’s domain organization with A: adenylation, T: thiolation, C: condensation domains. Numbers indicate amino acid positions of A and C domain borders (T domain positions not shown). The total length is 4884 amino acids. The diagram was drawn to scale. Triangles indicate corresponding intron positions. The dark regions indicate the amino acid alignment with black (highly similar) to white (not similar) sequences with respect to the consensus sequence. The amino acid alignment below shows Ds: D. septosporum, Cf: C. fulvum (protein ID: 193954), Zc: Zasmidium cellare (protein ID: 53743) and Mf: Mycosphaerella fijiensis (protein ID: 83197). The coloured bar represents consensus identity (below 30%); green shades indicate higher levels of identity (up to 100%) with bar height proportional to % identity.
The predicted gene model for \textit{DsNps2} has two introns. The expression levels of \textit{DsNps2} were very low both in culture and \textit{in planta} (below 10 RPMK) (Figure 3.1). However, the numbers of RNA-seq reads were higher at the 3’ end of the gene compared to 5’ end, a trend we observed in RNA sequencing analyses of long genes in the transcriptome analysis. Therefore, it was possible to confirm the intron at 3’ end with cDNA analysis (Appendix 6). Since the nucleotide and protein sizes may differ significantly due to a misannotated intron, the domain sizes where the introns correspond may also give hints about the gene model. The A domain that contains the first intron has a similar size and sequence motifs with the other A domains in the gene model (Figure 3.9), which suggests the first intron is probably correct.

Therefore, supported by the alignments (Figure 3.9, Appendix 20b), domain sizes and RNA-seq analyses (Appendix 6), the current gene model for \textit{DsNps2} appears correct. Further confirmation of the gene model could be done by cDNA sequencing of \textit{DsNps2} following use of growth conditions or a mutant that overexpresses \textit{DsNps2}.

\textbf{3.2.2.2. Comparison of \textit{DsNps2} from 19 strains of \textit{D. septosporum}}

\textit{DsNps2} gene sequences across the 19 strains were compared by aligning the sequences and determining the evolutionary selection pressure based on the \textit{dN/dS} ratio (Figure 3.10). The \textit{dN/dS} ratio of \textit{DsNps2} is 0.27, indicating negative selection pressure on the gene. The alignment also showed that the stop codon is absent or mutated in four strains, COLN, COLS, GUA1 and GUA2. However, the presence of another stop codon after 6 nucleotides means that the DsNps2 protein is only 2 amino acid longer in these four strains compared to that of the other strains. No positively selected codons were found for \textit{DsNps2}. 
Figure 3.10. *DsNps2* nucleotide alignment of NZE10 with 18 additional strains. Red arrows show the domain positions with A: Adenylation, T: thiolation, C: condensation. Yellow arrows indicate intron positions. The coloured bar represents consensus identity with brown indicating identities above 30% and below 100% and green indicating 100% identity. For each sequence, vertical bars on the top and bottom lines show differences in nucleotide and amino acid sequences compared to NZE10. A bar on the top line alone shows a synonymous mutation and a bar on both lines shows a non-synonymous mutation. Strains with stop-codon mutations are indicated with asterisk. Strains used in this analysis are listed in Appendix 5.
3.2.2.3. Gene cluster analysis

DsNps2 putative gene cluster analysis was done by searching 20 genes upstream and downstream of DsNps2, as outlined in Section 2.14.5. Because DsNps2 is the ortholog of C. fulvum Nps2 (CfNps2) (de Wit et al., 2012), it was also possible to compare the CfNps2 predicted gene cluster (Collemare et al., 2014) with that of DsNps2 in order to observe the synteny between them (Figure 3.11). However, there were differences in the methods used in the prediction of gene clusters between C. fulvum and D. septosporum putative SM gene clusters. The borders of C. fulvum SM gene clusters were defined when three successive genes did not have a predicted function related with secondary metabolism, or when two of the annotated genes had more than 5 kb between them (Collemare et al., 2014). Therefore, there are differences in C. fulvum SM putative gene clusters between this project and Collemare et al. (2014). Table 3.4 shows the predicted gene products for the genes in the DsNps2 cluster. The predicted gene cluster of DsNps2 is composed of 5 genes in a 40 kb region of chromosome 8. The biggest distance between two genes belonging to the predicted gene cluster is 12.6 kb, which contains 5 genes with no known secondary metabolism-related function. The gene clusters of DsNps2 and CfNps2 were found to be syntenic with the only exception that DsNps2 has one additional interfering genes compared to CfNps2.

Although the expression of DsNps2 was very low (<10 RPMK) (Section 3.1.2), Figure 3.11 suggests that all genes except Ds73932 were expressed significantly higher at late stage compared to early stage; of these, Ds55453 showed a very similar expression pattern to DsNps2. The gene that showed a different expression pattern, Ds73932, was expressed significantly higher in early stage than mid and late stages of infection. Since Ds73932 is predicted to encode a transcription factor, it is possible that this gene is responsible for regulating the expression of other genes within the predicted gene cluster. Similar to DsNps2, expression of CfNps2 was reported to be very low in culture and in planta at all stages of infection (Mesarich et al., 2014).
Figure 3.11. Synteny of DsNps2 and C. fulvum Nps2 putative gene clusters and expression profile of DsNps2 putative gene cluster in planta. Numbers inside arrows indicate JGI protein IDs. Pink: Secondary metabolite core genes DsNps2 and C. fulvum Nps2. Purple: putative DsNps2 and C. fulvum Nps2 cluster genes. Small white arrows: genes without predicted SM functions. Significant differences between gene expression in early-mid and mid-late stages of plant infection are marked with “*” and “+”, respectively (p<0.05). The differences between early and late stages were also significant for all five genes (p<0.05).

Table 3.4. Genes within DsNps2 putative gene cluster.

<table>
<thead>
<tr>
<th>JGI Protein ID</th>
<th>Predicted protein</th>
<th>Best hit organism/JGI ID</th>
<th>E / identity %</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Ds81928</td>
<td>Major facilitator transporter</td>
<td>Cladosporium fulvum/193959</td>
<td>3.8e-157/80.7</td>
</tr>
<tr>
<td>*Ds73932</td>
<td>Zinc-finger transcription factor</td>
<td>Cladosporium fulvum/193955</td>
<td>1e-54/63.5</td>
</tr>
<tr>
<td>*Ds90481/Nps2</td>
<td>Non-ribosomal peptide synthetase</td>
<td>Cladosporium fulvum/193954</td>
<td>0/76.6</td>
</tr>
<tr>
<td>*Ds55453</td>
<td>ABC transporter related protein</td>
<td>Cladosporium fulvum/193953</td>
<td>0/79.7</td>
</tr>
<tr>
<td>*Ds73935</td>
<td>Lysine/ornithine N-monoxygenase</td>
<td>Cladosporium fulvum/193952</td>
<td>0/90</td>
</tr>
</tbody>
</table>

*a http://genome.jgi.doe.gov/Dotse1/Dotse1.home.html
*b Predicted proteins according to antiSMASH and InterProScan analyses
*Asterisk indicates that genes with similar predicted functions present in other fungal NPS gene clusters.
*c All BlastP hits were confirmed true with reciprocal blast analyses.
DsNps2 was predicted to function in siderophore biosynthesis (Table 3.1), and previously predicted to be responsible for the synthesis of a ferricrocin-type siderophore (Collemare et al., 2014). Ferricrocins are responsible for iron storage and trans-cellular iron transport (Wallner et al., 2009). Although their primary roles are internal storage, secretion of small amounts of ferricrocin was reported in A. fumigatus (Hissen et al., 2004). The putative gene cluster of DsNps2 contains two transporter genes, ABC and MFS type (Table 3.4). Genes with similar functions can also be found in other fungal siderophore gene clusters (Schrettl et al., 2008; Haas et al., 2008; Philpott and Protchenko; 2008). In the fungal pathogen Histoplasma capsulatum, both ABC and MFS type transporters are present in a single siderophore gene cluster that is induced by low iron levels, and all genes except the one encoding the MFS type transporter are adjacent in a 25 kb region (Hwang et al., 2008). Therefore, if the prediction for DsNps2 is correct and the Nps2 metabolite is a ferricrocin, whether the transporters Ds81928 and Ds55453 are important functional components of the gene cluster is not known. However, due to its close proximity to the core gene DsNps2, the gene encoding for ABC transporter, Ds55453, is probably part of the putative gene cluster.

In addition to the transporters and the core gene, the DsNps2 putative gene cluster contains genes predicted to encode a zinc-finger transcription factor and a lysine/ornithine N-monooxygenase. Genes encoding proteins with similar predicted functions were reported in other fungal siderophore gene clusters. One gene, urbs1, predicted to encode a zinc-finger transcription factor was found in the fungus that causes corn smut disease, Ustilago maydis. Disruption of this gene leads to constitutive expression of the first gene in the siderophore biosynthesis pathway (sid1), and unregulated siderophore ferrichrome production (Yuan et al., 2001). The next gene in the putative DsNps2 gene cluster encodes a lysine/ornithine N-monooxygenase. Similar genes are present in the siderophore biosynthetic gene clusters of Omphalotus olearius, Histoplasma capsulatum and Ustilago maydis (Welzel et al., 2005, Hwang et al., 2008, Yuan et al., 2001). However, it is possible there are additional genes involved in the biosynthesis of the siderophore in D. septosporum. As mentioned before, the SM produced by the DsNps2 gene cluster may be a ferricrocin (Collemare et al., 2014). Like most fungal
siderophores, ferricrocins belong to the hydroxamate class of siderophores (Pourhassan et al., 2014). Biosynthesis of hydroxamates involves combined activity of lysine/ornithine N-monooxygenase and acyltransferase (Welzel et al., 2005), and thus an acyltransferase gene may also be involved in the biosynthesis of the Nps2 SM. However, a gene predicted to encode an acyltransferase was not found among 20 genes upstream and downstream of DsNps2.

To summarize, the gene model for DsNps2 is correct based on the alignment with closest homologs and cDNA reads. DsNps2 has an ortholog, CfNps2, and is under negative selection pressure. The predicted gene cluster for DsNps2 is composed of 5 genes that are conserved in other siderophore gene clusters. However, if the DsNps2 SM is a ferricrocin, it is possible that the transporter genes are not functional components of the putative DsNps2 gene cluster and an additional acetyltransferase gene may be involved in its synthesis.

### 3.2.3. NPS3

As explained in Section 3.1.2, DsNps3 was the most highly expressed D. septosporum SM core gene at the early stage of plant infection (Figure 3.1). Therefore, DsNps3 was of particular interest and additional analyses were done on this gene, including metabolite production and phenotypic analysis of a DsNps3 knockout mutant (ΔDsNps3) and phylogenetic analysis.

#### 3.2.3.1. Phylogenetic analysis

Phylogenetic analysis was done in order to find out if there are any functionally characterised orthologs of DsNps3 as preliminary FASTA analyses didn’t reveal any similar proteins. Phylogenetic analysis is a more accurate method for detecting homologous proteins than BlastP and FASTA (Koski and Golding, 2001). The analysis was performed by aligning the best reciprocal FASTA hits from each species belonging to the classes Dothideomycetes, Eurotiomycetes, and
Sordariomycetes represented in the JGI database, and building a phylogenetic tree using the maximum likelihood (ML) method as outlined in Section 2.14.3.

Figure 3.12. Phylogenetic tree and alignment of the peptide sequences of DsNps3 and its best FASTA hits. Red square shows *D. septosporum* (Dotse1) and its closest homologs. Dothideomycetes are indicated with asterisks. Numbers on nodes indicate approximate likelihood ratio test (aLRT) scores. Grey bars and gaps next to the species names show the alignment of the sequences used in the phylogenetic analysis. Details of the genome sequences used are presented in Appendix 23.
Initial phylogenetic analysis of DsNps3 (Figure 3.12) suggested that only the proteins within the red square are putative orthologs, while the other proteins are probably less related to DsNps3. *Cenococcum geophilum* (Cenge1/3) proteins were not included in the putative orthologs due to their low aLRT scores (0.088). Therefore, in order to minimise the gaps seen in the alignments, and obtain a DsNps3-focused phylogenetic tree, a new phylogenetic tree and alignment containing only the sequences within the red square were generated to focus on putative DsNps3 orthologs (Figure 3.13).

**Figure 3.13. Phylogenetic tree and alignment of the amino acid sequences of putative DsNps3 orthologs.** a) Phylogenetic tree of DsNps3 with its putative orthologs. Red (Dotse1): DsNps3. Numbers on branches indicate approximate likelihood ratio test (aLRT) scores. b) Alignment of the sequences used in the phylogenetic analysis. Names on the left show the abbreviation of species names as in JGI. The coloured bar represents consensus identity with red (below 30%); green shades indicate higher levels of identity (up to 100%) with bar height proportional to % identity.
The phylogenetic tree and amino acid alignment presented in Figure 3.13 showed that most of the gaps between the protein sequences of DsNps3 and its putative orthologs in the initial alignment (Figure 3.12; red box) were due to the presence of the other less similar sequences. The most similar proteins to DsNps3 were from the Sordariomycete fungi *Colletotrichum somersetensis* (Colso1) (protein ID: 75126), and *M. grisea* (Maggr1) (protein ID: 116173), and the Dothideomycete *C. fulvum* (Clafu1) (protein ID: 185841, CfNps1). A big gap in the first half of CfNps1 was seen in the alignment, supporting the hypothesis of Collemare et al. (2014) that the *CfNps1* gene is truncated. The C-terminus of DsNps3 was shorter than that of its putative orthologs in the alignment. One possible reason for this was a frameshift or a nonsense mutation at the 3' end of *DsNps3*, but a three-frame translation of this region did not reveal any similarity to the extended C termini of the putative orthologs. Since all of the putative orthologs of DsNps3 are in Sordariomycetes fungi except the truncated CfNps1, DsNps3 may be unique in structure among the Dothideomycetes.

Of the proteins in Figure 3.13 only a few have been functionally characterised. The SM produced by FOXG_11847 (named beas after functional analysis) of *Fusarium oxysporum* was confirmed to be a virulence factor for both mice and tomato plants (López-Berges et al., 2013). SMs produced by the orthologs of *beas*, enniatin synthase from *F. avenaceum* and beauvericin synthase from *Beauveria bassiana* (López-Berges et al., 2013), were functionally characterised virulence factors on their respective potato and insect hosts (Xu et al., 2008; Herrmann et al., 1996). Therefore, there are precedents for orthologs of *DsNps3* as virulence factors.

### 3.2.3.2. Confirmation of gene model for *DsNps3*

Gene model confirmation was done by analysing the alignment of the predicted DsNps3 amino acid sequence with those of most closely related protein sequences and in culture and *in planta* RNA-seq data.

Figure 3.14 illustrates the domain organization of the DsNps3 protein (protein ID: 71189) and its alignment with its closest homologs according to the phylogenetic
analysis shown in Section 3.2.3.1. The protein alignment of DsNps3 has low overall consensus and DsNps3 amino acid identities to *C. fulvum*, *C. somersetensis*, and *M. grisea* were 42.72% (E=0), 42.59% (E=0), and 43.97% (E=0), respectively. The alignment also included at least three big gaps. However, none of the big gaps were in the domain positions and domains were more conserved than the rest of the alignment. Therefore, even though the alignment of DsNps3 with its closest homologs showed no evidence that the JGI gene model for *DsNps3* is incorrect, no clear decision could be made on whether the *DsNps3* gene model is correct due to low similarity between DsNps3 and its closest homologs.

The predicted gene model for *DsNps3* has no introns and this was confirmed by continuous cDNA reads from start codon to stop codon in the RNA-seq data that were mapped to the JGI gene model for *DsNps3* (Appendix 7). Therefore, although the alignment of DsNps3 with its closest homologs does not provide evidence for whether the JGI gene model is correct because of the low similarity of DsNps3 with its closest homologs, it is supported by the RNA-seq data.
Figure 3.14. Domain structure of DsNps3 and its protein alignment with closest homologs. DsNps3 indicates the peptide’s domain organization with A: adenylation, T: thiolation, C: condensation, NM: N-methyltransferase domains. Numbers indicate amino acid positions of domain borders (borders of domains that are too close to another border are not shown). Dotted lines show positions of the DsNps3 gaps in the alignment. The dark regions indicate the amino acid alignment with black (highly similar) to white (not similar) sequences with respect to the consensus sequence. The total length is 2611 amino acids. The diagram was drawn to scale. The amino acid alignment below shows Ds: D. septosporum, Cf: Cladosporium fulvum (protein ID: 185841), Cs: C. somersetensis (protein ID: 75126), Mg: Magnaporthe grisea (protein ID: 116173). The coloured bar represents consensus identity with red (below 30%); green shades indicate higher levels of identity (up to 100%) with bar height proportional to % identity.
3.2.3.1. Comparison of DsNps3 from 19 strains of D. septosporum

Comparison of DsNps3 across 19 D. septosporum strains was done by aligning the DsNps3 nucleotide sequences and determining the evolutionary selection pressure. Figure 3.15 shows DsNps3 nucleotide sequence alignment of NZE10 with 18 additional D. septosporum genomes. For DsNps3, dN/dS is 0.27, therefore, DsNps3 is under negative selection, suggesting that the gene might have a critical role for D. septosporum. However, the alignment presented in Figure 3.15 shows that four codons, including one codon from the first A domain and one codon from the NM domain are under statistically significant (P>0.95) positive selection. This alignment also supported the gene model analysis discussed in Section 3.2.3.2, suggesting that the gene model is correct. It was not possible to determine negatively selected sites due to lack of DsNps3 orthologs in other species and the highly similar sequences for the 19 D. septosporum genomes.

In conclusion, according to the dN/dS ratio DsNps3 is under negative selection but has four statistically significant positively selected sites. Alignment of DsNps3 from the 19 D. septosporum genomes also supports the JGI gene model for DsNps3.
Figure 3.15. *DsNps3* nucleotide alignment of NZE10 with 18 additional genomes. Red arrows show the domain positions with A: adenylation, T: thiolation, C: condensation, NM: N-methylation. Bold black lines above domains show statistically significant positively selected codons (P>0.95). Nucleotide sequence of *DsNps3* is 7833 bp. The coloured bar represents consensus identity with brown indicating identities above 30% and below 100% and green indicating 100% identity. For each sequence, vertical bars on the top and bottom lines show differences in nucleotide and amino acid sequences compared to NZE10. A bar on the top line alone shows a synonymous mutation and a bar on both lines shows a non-synonymous mutation. Strains used in this analysis are listed in Appendix 5.
3.2.3.2. Gene cluster analysis

Gene cluster analysis was performed on DsNps3 and its neighbouring genes as explained in Section 2.14.5 in order to predict the genes that might take part in the production of the corresponding nonribosomal peptide. Figure 3.16 shows the putative gene cluster for DsNps3, possible regulatory genes of the gene cluster and expression of genes in the putative gene cluster and the possible regulatory genes. Predicted functions of the genes in the DsNps3 putative gene cluster are shown in Figure 3.16. Since DsNps3 was a gene of main interest in this project, genes with potential regulatory roles were included in the gene cluster analysis. The DsNps3 putative gene cluster spans a 44 kb region containing 4 genes in chromosome 4. In this region, there are also 2 potential regulatory genes and 6 genes with no predicted functions related to secondary metabolism. The biggest distance between two genes belonging to the putative gene cluster is 24 kb, containing 1 potential regulatory gene and 6 genes with no SM-related functions.

Gene cluster analysis of DsNps3 (Figure 3.16) showed that none of the genes in the putative gene cluster were co-expressed with DsNps3, which was expressed significantly higher at early stage compared to mid and late stages. Ds71190 and Ds170752 were expressed significantly higher at mid-stage compared to early and late stages, and Ds71198 was most highly expressed at late stage of infection. Genes within the SM gene clusters are generally co-expressed (Keller and Hohn, 1997), however, this is not always the case (Turgeon and Bushley, 2010). An example to a SM gene cluster that is not co-expressed is the dothistromin gene cluster of D. septosporum, explained in detail in Section 1.4.5.1.
Figure 3.16. Organization and expression of \textit{DsNps3} and its neighbouring genes in early, mid, and late stages of infection. Numbers inside arrows indicate JGI protein IDs. Pink: Secondary metabolite core gene (\textit{DsNps3}). Purple: putative \textit{DsNps3} cluster genes. Yellow: Genes that are not part of the putative gene cluster but have potential regulatory roles. Small white arrows: genes without predicted SM functions. Significant differences between early-mid and mid-late stages were marked with “*” and “+”, respectively (p<0.05). The differences between early and late stages were also significant for all cluster and potential regulatory genes except Ds71191.

**Table 3.5.** Genes within \textit{DsNps3} putative gene cluster.

<table>
<thead>
<tr>
<th>JGI Protein ID</th>
<th>Predicted protein</th>
<th>Best hit organism/JGI ID</th>
<th>E / identity %</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Ds71189/Nps3</td>
<td>Non-ribosomal peptide synthetase</td>
<td>\textit{Penicillium expansum}/371848</td>
<td>8.5e-187/52.4</td>
</tr>
<tr>
<td>*Ds71190</td>
<td>Alcohol dehydrogenase, class V</td>
<td>\textit{Cladosporium fulvum}/189228</td>
<td>0/95.8</td>
</tr>
<tr>
<td>*Ds719752</td>
<td>Short chain dehydrogenase/reductase</td>
<td>\textit{Mycospherella fijiensis}/57497</td>
<td>1.2e-103/64.2</td>
</tr>
<tr>
<td>*Ds71198</td>
<td>Flavin-containing monooxygenase</td>
<td>\textit{Cladosporium fulvum}/196698</td>
<td>0/93.6</td>
</tr>
</tbody>
</table>

\(^a\) [http://genome.jgi.doe.gov/Dotse1/Dotse1.home.html](http://genome.jgi.doe.gov/Dotse1/Dotse1.home.html)
\(^b\) Predicted proteins according to antiSMASH and InterProScan

Asterisk indicates genes with similar predicted functions present in other fungal NPS gene clusters.

* Genes with similar predicted function to Ds71190 was found in other fungal SM gene clusters such as dothistromin gene cluster (Bradshaw et al., 2013).

All BlastP hits were confirmed true with reciprocal blast analyses.
The predicted functions of two of the three genes in the putative DsNps3 gene cluster are similar to those predicted in other NRPS gene clusters. A gene encoding for a short chain dehydrogenase/reductase is present in the NRPS gene cluster of apicidin belonging to F. incarnatum (Jin et al., 2010), and there is a gene predicted to encode for a short chain dehydrogenase 25 kb downstream of DsNps3. In addition, monooxygenases are another common element found in SM gene clusters (Zhang et al., 2014; Collemare et al., 2014) and a predicted flavin containing monooxygenase gene is also present in the putative gene cluster of DsNps3.

The third gene (Ds71190), just downstream of DsNps3, is predicted to encode an alcohol dehydrogenase. Such genes are commonly found in fungal indole diterpene or PKS gene clusters such as the dothistromin gene cluster of D. septosporum (Young et al., 2001, Bradshaw et al., 2013), but have not been reported in NRPS gene clusters.

Aside from the biosynthetic genes belonging to the DsNps3 putative gene cluster, two genes that may have a regulatory role were found. Genes Ds71191 and Ds71196 are predicted to encode a splicing coactivator and a sensory transduction histidine kinase, respectively. Studies on a splicing coactivator gene from A. carbonarius showed a high correlation between the expression of a possible regulatory splicing coactivator gene and ochratoxin A production (Lunardi et al., 2009). A gene encoding an osmotic signal transduction histidine kinase was reported to take part in the regulation of secondary metabolism of Fusarium graminearum (Ochiai et al., 2007).

The predicted DsNps3 SM was cyclosporin according to NaPDoS prediction (Table 3.1). Gene clusters responsible for cyclosporin production were explained in detail at Section 3.2.1.3. However, genes in the DsNps3 putative gene cluster do not support this prediction.
3.2.3.3. Knockout of DsNps3

In order to perform functional analysis of the DsNps3 gene, deletion mutants were prepared. However, because DsNps3 is a very large gene, only two of the essential domain regions were deleted and a reading frame shift was ensured. First, the DsNps3-pOSCAR KO construct (Appendix 8) was prepared as explained in Section 2.7.2. Following transformation of the KO construct into D. septosporum NZE10 protoplasts as outlined in Section 2.9, 38 hygromycin-resistant colonies were obtained. From them, two single-spore purified colonies, derived from one transformant, were confirmed to be ΔDsNps3 strains by PCR, using first a primer pair that only amplifies wild type but not ΔDsNps3 (Figure 3.18a), and then using two primer pairs that each targeted the hygromycin coding region and 5' or 3' flanking regions (Figure 3.18b). Primer binding positions are shown in Figure 3.17.

![Figure 3.17 Schematic representation of wild type and KO mutants of DsNps3.](image)

The positions of EcoRV restriction enzyme recognition sites, primer-binding sites, flanking regions, knockout target region in wild type (A and T domains) and hygromycin resistance gene (hph) with TrpC promoter were shown.
Figure 3.18. PCR-based confirmation of purified colonies for the absence of WT genotype. 12 colonies were sub-cultured from a single-spore purified colony were analysed using PCR-based confirmation that amplify wild type but not $\Delta$DsNps3 to confirm successful purification using primers NPS3-FP-wtcheck and D. septo NRPS 3 reverse (325 bp). Wt: Wild type. NTC: No template control.

Figure 3.19. PCR-based confirmation of $\Delta$DsNps3 candidates for the presence of the KO genotype. The colonies 6 and 9 that were identified lacking WT genotype (named KO1 and KO2, respectively) were analysed using PCR amplification with primers Nps3-5FP-Screen and OscHygR for 5' end (1793 bp) and OscHygF and NPS3-3RP-domainscreen for 3' end (1414 bp).

Lack of PCR products in colonies 6 and 9 in the PCR amplifications presented in Figure 3.18a suggests the loss of DsNps3 by gene replacement in these two colonies. Therefore, the colonies 6 and 9 (named as KO1 and KO2, respectively), which were purified from the same transformant, were confirmed not to contain
any wild type gDNA, confirming successful purification. PCR amplification presented in Figure 3.18b confirmed the ΔDsNps3 mutants by amplifying 5’ and 3’ ends of the hygromycin coding region and the flanking sites.

Following preliminary screening of ΔDsNps3 candidates, Southern blot hybridization was used to confirm single copy integration of the KO construct at the correct position in the ΔDsNps3 candidates, KO1 and KO2 (Appendix 9).

Confirming ΔDsNps3 using PCR and Southern hybridization enabled future analyses for the determination of the metabolite and the effects of DsNps3 gene loss for D. septosporum.

3.2.3.4. TLC analysis

In order to determine the Nps3 SM, TLC analysis was performed by comparing the SMs extracted from wild type and ΔDsNps3 strains. Preliminary TLC analyses were done to determine optimum conditions for Nps3 SM extraction by using organic solvents with increasing polarity each time, based on suggestions of Dr. Mark Patchett (Massey University) (Figure 3.20). Preliminary screenings based on these solvents showed differences in the TLC band profiles of wild type and ΔDsNps3 (Figure 3.20) in the culture filtrate. No difference between wild type and ΔDsNps3 was observed in the metabolites extracted from mycelia.
Figure 3.19. Solvent optimization for Nps3 SM detection in TLC analysis. Numbers and letters indicate the solvent used for the extraction of metabolites. F1: EtAc, F2-A: Chloroform, F2-B: Dichloromethane after chloroform, F3-A: Dichloromethane, F3-B: 1-Butanol after dichloromethane. wt: Wild type, KO1 and KO2: ΔDsNps3 D. septosporum strains. The metabolite present in wild type but missing in ΔDsNps3 was shown with an asterisk.

Initial TLC analysis for the solvent optimization suggested that the metabolites extracted using chloroform and dichloromethane showed differences in the TLC profiles between wild type and ΔDsNps3 strains (Figure 3.20). The metabolite represented with an asterisk was initially thought a candidate for Nps3 SM. Therefore, dichloromethane was used for extraction for Nps3 metabolite determination.
For comparison of the metabolites extracted from $\Delta DsNps3$ and wild type $D. septosporum$ strains, the amounts of metabolites spotted onto the TLC plate were normalized based on dothistromin levels in the samples estimated by measuring intensity of TLC bands (Section 2.11.2). In addition, the amount of metabolite dried down before resuspension and spotting onto the TLC plate for each strain was increased from 200 μL for each strain to 10 mL, 5.9 mL and 6.4 mL for wild type, KO1 and KO2 respectively, enabling visualization of faint bands. Figure 3.21 shows the TLC analysis of wild type and $\Delta DsNps3$ using approximately equal amount of metabolites. This showed that the band that was initially thought missing in $\Delta DsNps3$ (Figure 3.20) was present in both wild type and $\Delta DsNps3$ strains. However, another band that was too faint to visualise in the initial TLC screening (Figure 3.20) was present in wild type and missing in both $\Delta DsNps3$ strains.

The TLC profile analysis presented in (Figure 3.21) shows one band present in the wild type and missing in the $\Delta DsNps3$ KO1 and KO2 strains that might be a candidate for the SM synthesized by Nps3. However, the result was not very clear, as there were multiple faint bands very near to each other. There was also another difference in the band profiles between wild type and $\Delta DsNps3$ KO1 and KO2 strains that was not visible to naked eye, but detected by a small peak in the quantification analysis. These two bands may represent Nps3 SM or its intermediates. Another possibility is that Nps3 SM might not be produced at high levels in PMMG. These problems might be solved by changing the growth conditions and applying better extraction and separation methods. It should also be noted that no difference was observed on TLC profiles of wt and Nps3 KO mutants if the metabolites were extracted with ethyl acetate, the solvent used for extracting metabolites in Section 3.1.3. Therefore, this band was not observed at all on analysis of the effects of different media conditions.

To sum up, in the TLC analysis of metabolites extracted from wild type and $\Delta DsNps3$ using dichloromethane, a candidate for Nps3 SM was present in wild type but missing in the $\Delta DsNps3$. However, since the candidate metabolite was not produced at high levels and not separated sufficiently, no chemical characterization was performed.
Figure 3.20. TLC picture and band profiles of *D. septosporum* wild type and ΔDsNps3 strains. Metabolites that were present in wild type and missing in ΔDsNps3 samples is indicated with asterisks. Dothistromin was used as control and the peak corresponding dothistromin band is indicated with the letter D.
3.2.3.5. Sporulation and growth rate analyses

Fungal secondary metabolites such as aflatoxin and sterigmatocystin are usually associated with asexual sporulation, and there are many SMs reported to affect sporulation (Skory et al., 1993; Calvo et al., 2002). Both sporulation and metabolite biosynthesis are usually negatively regulated by G-protein-mediated signalling pathways when the growth signals are down (Hicks et al., 1997). On the other hand, loss of some secondary metabolites cause reduced growth rate in fungi (Calvo et al., 2002). Therefore, to test the hypothesis that DsNps3 affects sporulation, wild type and ΔDsNps3 spores were counted as outlined in Section 2.12.2. Radial growth rate analysis was also done by measuring wild type and ΔDsNps3 colony diameter every 3 days over a period of 30 days (Section 2.12.3).

Figure 3.21. Sporulation and radial growth rates of D. septosporum wild type and ΔDsNps3 strains. a) Number of spores/mL of culture x10⁷ for wild type and ΔDsNps3 KO2 strains. Asterisk shows the significant difference (p<0.05) between wild type and ΔDsNps3 KO2. Two technical replicates were counted for each of three biological replicates. b) Average growth rates (mm/day) of wild type and ΔDsNps3 KO1 and KO2 strains. Four biological replicates were counted.
The number of spores was significantly lower (p<0.05) by 33% in ΔDsNps3 compared to wild type *D. septosporum* (Figure 3.22a). ΔDsNps3 KO1 was not included in the sporulation rate analysis because of dense mycelia that interfered with spore counting that was most likely because of failure to apply equal pressure on the plates for release of spores, which resulted in stripping of mycelia from the KO6 plate. Radial growth rate analysis results showed no significant difference (p<0.05) between wild type and ΔDsNps3 KO strains (n=8) (Figure 3.22b).

In conclusion, *D. septosporum* ΔDsNps3 showed significantly lower sporulation compared to wild type, but no difference in radial growth rate.

### 3.2.3.6. Pathogenicity assay

To determine if the SM produced by the *DsNps3* gene cluster is a virulence factor, pine seedlings were inoculated with wild type or ΔDsNps3 *D. septosporum* spores. Comparisons were made of the fungal surface network at an early stage of infection. Then the timing of first lesion appearance, lesion size, number, and fungal biomass estimations between wild type and ΔDsNps3-infected needles were compared as outlined in Section 2.12.1.

The first step of plant fungal infection is the plant host surface interaction, and this step plays a vital role in the early stages of the disease process. At this stage, attachment to plant surface, spore germination, recognition of the host, infection structure formation and finally the penetration of the host tissue occur (Knogge et al., 1998). Since *DsNps3* was expressed at a significantly higher level at early stage compared to mid and late stages (Figure 3.1), it is possible that Nps3 SM might play a role in the early stages of the infection. Fungal surface network at early stage of *D. septosporum* infection was compared between wild type and ΔDsNps3 to test the hypothesis that DsNps3 has a role during the asymptomatic period of the infection. The percentage area covered by the *D. septosporum* infected pine needles was significantly less in ΔDsNps3 KO2 compared to wild type. Figure 3.23 shows representative images of the surface network analysis of wild type and ΔDsNps3.
All of the fluorescence microscopy images used in this analysis are presented in Appendix 10.

Figure 3.22 Hyphal network on pine needle surface infected with wild type and ΔDsNps3 *D. septosporum*. a) Sample fluorescence microscopy images of wild type and ΔDsNps3 infected P3 at 15 dpi. b) Percentage of needle surface covered by fungal hyphae. Different plant genotypes are named P1-P4. WT: Wild type, KO2: ΔDsNps3. Significant differences between wt and ΔDsNps3 strains (p<0.05) are shown with an asterisk. The largest three surface networks were measured per plant genotype for each strain.
Surface network analysis presented in Figure 3.23 show that the hyphal network on the needle surface was significantly lower in three of the four plant genotypes tested in ΔDsNps3 compared to wild type. KO1 was not included in the analysis since they are the same mutants from a single transformant (Section 3.2.3.3).

In order to test if Nps3 SM may be a virulence factor, the ratio of fungal to plant biomass in the lesions obtained from wild type and ΔDsNps3 infected needles was estimated using real-time PCR with known single-copy *D. septosporum* and *P. radiata* genes as explained in Section 2.12.1.2. Figure 3.24 shows the comparison of estimated fungus to plant biomass ratios of ΔDsNps3-infected lesions with those of wild type infected lesions.

![Figure 3.23. Fungal biomass ratio of ΔDsNps3-infected lesions to wild type *D. septosporum* infected lesions.](image)

The needles were harvested at 18 wpi.

The analysis in Figure 3.24 showed that fungus to plant biomass ratios of ΔDsNps3-infected lesions were less than 30% of that of wild type-infected lesions. The difference in ratios of wild type and ΔDsNps3-infected lesions of the plant genotype P1 was not significant due to variations up to 2 fold between replicates. Overall, early stage hyphal surface network and fungal biomass assay results suggest that the Nps3 SM may be a virulence factor. Full biomass estimation assay results are presented in Appendix 11.

Even though surface network significantly decreased in ΔDsNps3 compared to wild type, no significant differences were observed in the timing of first lesion
appearance, lesion size, or number between wild type and ΔDsNps3 (Table 3.6). Lesion pictures for each plant genotype and strain are shown in Appendix 12.

### Table 3.6 Lesion number, size and first lesion appearance in wild type and ΔDsNps3 D. septosporum infected pine needles.

<table>
<thead>
<tr>
<th></th>
<th>wt-P1</th>
<th>KO2:P1</th>
<th>wt-P2</th>
<th>KO2-P2</th>
<th>wt-P4</th>
<th>KO2-P4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lesion number</strong></td>
<td>38</td>
<td>8</td>
<td>11</td>
<td>7</td>
<td>36</td>
<td>17</td>
</tr>
<tr>
<td><strong>Lesion size (mm)</strong></td>
<td>2.1±1.3</td>
<td>1.8±1.4</td>
<td>1.5±0.9</td>
<td>1.5±0.6</td>
<td>2.2±1.2</td>
<td>2.4±1.1</td>
</tr>
<tr>
<td><strong>First lesion</strong></td>
<td>7</td>
<td>7</td>
<td>8+</td>
<td>8+</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

* Number of lesions per tree

A limitation with the pathogenicity assay was the difficulty of having successful infection. For example, no lesion was observed on the needles of the 4th plant genotype P3 even though there was a significant difference in the fungal surface network (Figure 3.23). It is challenging to achieve reliable and consistent D. septosporum infection using artificial inoculation under controlled conditions and high variations can occur in the duration infection stages (Kabir et al., 2013). Another possibility for observing no DNB symptoms in P3 could be due to heavy insect infestation on the P3 samples during the incubation period that might have increased basal defence responses of these plants. Also, the plants used in this experiment (P1-P4) had different genotypes, which contributed to the variations between different plants. Initial predictions on the resistance levels of the plants to the D. septosporum were P3>P4>P2>P1. The fact that the findings in this experiment didn't correlate with the initial predictions suggest that the impact of the insect infestation might have had a bigger role in the infection rates compared to that of plant genotypes.

To sum up, the predicted gene model for DsNps3 appears to be correct but phylogenetic analysis suggests that the Nps3 SM may be unique among Dothideomycetes. In addition, decreased sporulation rate, early stage fungal surface network, and fungal biomass in planta were observed with the ΔDsNps3 mutant compared to wild type. However, no difference was seen in radial growth rate, lesion number, size, or the timing of first lesion appearance in pathogenicity assays between ΔDsNps3 and wild type. Therefore, reduced fungal surface network...
in the early stage of infection and reduced fungal biomass in late stage lesions in ΔDsNps3 compared to WT suggest that the DsNps3 SM may be a virulence factor.
3.3. Polyketide synthases

*D. septosporum* has five predicted PKS core genes (Table 3.1). However, only *DsPks1*, *DsPks2*, and *DsPks3* were focused on in this project because the preliminary analyses showed that *DsPks4* is a truncated gene and *DsPksA* was extensively studied previously (Section 1.4.4.1).

3.3.1. PKS1

Among the *D. septosporum* SM core genes, *DsPks1* had the highest level of expression at the late stage of pine needle infection (Figure 3.1). Therefore, *DsPks1* was one of the key genes of interest as explained in Section 3.1.2 and phylogenetic analysis was done for DsPks1 in addition to the analyses done for all of the SM core genes.

3.3.1.1. Phylogenetic analysis

To determine if DsPks1 protein has any functionally characterised orthologs, phylogenetic analysis was done by aligning DsPks1 and its best reciprocal hits from all Dothideomycetes, Eurotiomycetes, and Sordariomycetes in the JGI database, followed by phylogenetic tree building using the Maximum Likelihood method as explained in Section 2.14.3. The phylogenetic tree of DsPks1 (Figure 3.25) showed that all of the closest homologs of DsPks1 were from fungi in the Dothideomycetes class. These proteins were from *Z. cellare* (protein ID: 19481), *C. fulvum* (protein ID: 191425), *M. graminicola* (protein ID: 96592), *M. fijiensis* (protein ID: 216850), *Cercospora zeae-maydis* (protein ID: 68161), and *Septoria musiva* (protein ID: 159092).

None of the putative orthologs of DsPks1 from the phylogenetic analysis presented in Figure 3.25 have been functionally characterised. However, CfPks1 is predicted to have a role in elsinochrome and/or melanin biosynthesis (Collemare et al., 2014). In addition, *M. graminicola* (now called *Zymoseptoria tritici*) PKS1 (protein
ID: 96592) is considered to be involved in DHN-melanin biosynthesis (Lendenmann et al., 2014) and the *C. lagenarium* ortholog of *MgPKS1* was functionally characterised to be involved in melanin biosynthesis using gene knockout and complementation assays (Kubo et al., 1991; Takano et al., 1995). Therefore, DsPks1 might be responsible for biosynthesis of a compound with similar structure to melanin or elsinochrome.

Fungal melanins have a variety of functions such as protection against exogeneous stress factors, contributing to cell wall related functions such as pressure-stabilization, and as virulence factors (reviewed by Scharf et al., 2014). For example, the rice blast fungus *M. oryzae* requires DHN-melanin to penetrate host cell walls during infection, as shown by the inability of melanin-deficient (albino) mutants to penetrate epidermal cells (Chumley and Valent, 1990). Elsinochrome is a non-host-selective toxin and virulence factor produced by *Elsinoë fawcettii* and *Elsinoë australis* (Chung, 2011). Therefore, considering the high expression level of *DsPks1* in planta and the virulence-associated roles of DHN-melanin in other fungi, it is possible that the DsPks1 SM may be a virulence factor.
Figure 3.24. Phylogenetic analysis of DsPks1. Yellow: *D. septosporum*. Red: Dothideomycetes. Blue: Eurotiomycetes. Green: Sordariomycetes. Numbers on the nodes indicate approximate likelihood ratio test (aLRT) scores. *D. septosporum* and its closest homologs are shown in a red square. Details of the genome sequences used are presented in Appendix 23.
3.3.1.2. Confirmation of gene model for \textit{DsPks1}

The JGI gene model for \textit{DsPks1} was confirmed by aligning its protein sequence with its closest homologs, including its \textit{C. fulvum} ortholog, \textit{CfPks1} (de Wit et al., 2012) as well as by analysing RNA-seq data. The DsPks1 protein alignment (Figure 3.26) has high overall consensus. Amino acid identities of DsPks1 with its closest homologs were 94.34\% (E=0) for CfPks1, 86.47\% (E=0) for the Pks of \textit{Zasmidium cellare}, and 79.67\% (E=0) for that of \textit{M. graminicola}, suggesting that the proteins within the alignments share similar functions. There was no gap or insertion in the DsPks1 sequence at the alignment, therefore supporting the JGI gene model for \textit{DsPks1}. The predicted gene model for \textit{DsPks1} has no introns and this was confirmed by the RNA-seq data (Appendix 13).

3.3.1.3. Comparison of \textit{DsPks1} from 19 strains of \textit{D. septosporum}

Evolutionary selection pressure on \textit{DsPks1} was analysed by aligning \textit{DsPks1} sequences from 19 \textit{D. septosporum} strains, from varying international origins, estimating the overall selection pressure using dN/dS ratio and detecting the sites that are under significant selection (Section 2.14.4). The dN/dS ratio of \textit{DsPks1} is the lowest of \textit{D. septosporum} SM core genes (dN/dS=0.13), indicating that \textit{DsPks1} is under strong negative selection. This suggests that the Pks1 SM, a predicted melanin, may be an important compound for \textit{D. septosporum}. Only one codon, just before the AT domain, showed significantly positive selection (P>0.99) (Figure 3.27).
Figure 3.25. Domain structure of DsPks1 and its protein alignment with closest homologs. DsPks1 indicates the peptide's domain organization with KS: keto-synthase, AT: acyltransferase, ACP: acyl carrier protein, TE: thioesterase domains. Numbers indicate amino acid positions of domain borders in addition to start of the DsPks1 peptide sequence. The total length is 2189 amino acids. The diagram was drawn to scale. The dark regions indicate the amino acid alignment with black (highly similar) to white (not similar) sequences with respect to the consensus sequence. The amino acid alignment shows Ds: D. septosporum Pks1, Cf: C. fulvum (protein ID: 191425), Zc: Z. cellare (protein ID: 19481), Mg: M. graminicola (protein ID: 96592). The coloured bar represents consensus identity with red (below 30%); green shades indicate higher levels of identity (up to 100%) with bar height proportional to % identity.
Figure 3.26. *DsPks1* nucleotide alignment of NZE10 with 18 additional strains. Red arrows show the domain positions with KS: keto-synthase, AT: acyltransferase, ACP: acyl carrier protein, TE: thioesterase. The bold black line just before the AT domain indicates the statistically significant positively selected codon (P>0.99). The coloured bar represents consensus identity with green shades indicating identities above 30% and below 100% and green indicating 100% identity. For each sequence, vertical bars on the top and bottom lines show differences in nucleotide and amino acid sequences compared to NZE10. A bar on the top line alone shows a synonymous mutation and a bar on both lines shows a non-synonymous mutation. Strains used in this analysis are listed in Appendix 5.
3.3.1.4. Gene cluster analysis

Gene cluster analysis was assisted by the knowledge that DsPks1 was predicted to function in melanin/elsinochrome biosynthesis (Table 3.1), and its ortholog CfPks1 (de Wit et al., 2012) has a predicted gene cluster (Collemare et. al. 2014). The DsPks1 putative gene cluster (Figure 3.28) contains 4 genes in a 27.4 kb region of chromosome 10. The biggest distance between two genes in the DsPks1 gene cluster is 10.8 kb, containing 3 genes with no predicted SM-related function. Synteny analysis between the putative gene clusters of DsPks1 and CfPks1 showed an inversion of Ds74621 and Ds178588 in the DsPks1 putative gene cluster (Figure 3.28). All genes except Ds37856 within the DsPks1 putative gene cluster were expressed at significantly higher levels at late stage compared to early and mid stages (Figure 3.28).
Figure 3.27. Synteny of DsPks1 and CfPks1 putative gene clusters and expression profile of DsPks1 putative gene cluster in planta. Numbers inside arrows indicate JGI protein IDs. Pink: Secondary metabolite core genes (DsPks1 and CfPks1). Purple: putative DsPks1 and CfPks1 cluster genes. Small white arrows: genes without predicted SM functions. Significant differences in the gene expression between early-mid and mid-late stages of infection were marked with "**" and "*", respectively (p<0.05). The differences between early and late stages were significant for all four genes (p<0.05).

<table>
<thead>
<tr>
<th>JGI Protein ID</th>
<th>Predicted protein</th>
<th>Best hit organism/JGI ID</th>
<th>E / identity %</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Ds74621</td>
<td>Zinc-finger transcription factor</td>
<td>Cladosporium fulvum /191427</td>
<td>0/91.5</td>
</tr>
<tr>
<td>*Ds178588/Bm1</td>
<td>Short-chain dehydrogenase/reductase</td>
<td>Cladosporium fulvum /191428</td>
<td>0/97.4</td>
</tr>
<tr>
<td>*Ds47338/Pks1</td>
<td>Polyketide synthase/Pks1</td>
<td>Cladosporium fulvum /191425</td>
<td>0/94.3</td>
</tr>
<tr>
<td>*Ds37856</td>
<td>Prefoldin</td>
<td>Cladosporium fulvum /191424</td>
<td>3.4e-119/91.8</td>
</tr>
</tbody>
</table>

* http://genome.jgi.doe.gov/Dotse1/Dotse1.home.html

* Predicted protein functions according to antiSMASH and InterProScan

* Indicate genes with similar predicted functions present in other fungal PKS gene clusters. All BlastP hits were confirmed true with reciprocal blast analyses.

Table 3.7. Genes within the DsPks1 putative gene cluster.

As mentioned previously, DsPks1 is predicted to be responsible for melanin or elsinochrome biosynthesis. Even though the metabolite produced by CfPks1 is not known, it is associated with a black pigment (Ökmen et al., 2014), whilst in contrast elsinochrome is a red/orange pigment (Liao and Chung, 2008). In addition, elsinochrome production was not detected in C. fulvum metabolite extracts, even though CfPks1 was expressed (Collemare et al., 2014). These suggest that the CfPks1 SM may be involved in production of melanin. Given the similarity of DsPks1 to CfPks1, DsPks1 SM also most likely produces melanin. The genes

121
within the *DsPks1* putative gene cluster have similar predicted functions to genes in other known fungal melanin/elsinochrome gene clusters.

The first gene in the *DsPks1* putative gene cluster is predicted to encode a zinc-finger transcription factor (Table 3.7). Zinc-finger transcription factors are regulatory elements common in secondary metabolite gene clusters, including *Amr1*, which positively regulates melanin biosynthesis in *Alternaria brassicicola* (Cho et al., 2012). The second gene, *Brn1*, is predicted to encode a short-chain dehydrogenase/reductase similar to those found in the melanin gene clusters of other fungi including *M. grisea* (Vidal-Cros et al., 1994) and *C. heterostrophus* (Eliahu et al., 2007). The final gene is predicted to encode prefoldin, a chaperone responsible for the transfer of unfolded proteins to cytosolic chaperonin (Vainberg et al., 1998). A gene predicted to encode prefoldin, *PRF1*, is present in the elsinochrome gene cluster of *Elsinoë fawcettii* (Chung, 2011). It was hypothesized that *PRF1* stabilizes some of the enzymes required for elsinochrome biosynthesis. Another gene predicted to encode a prefoldin, *MgPRF1*, is in the putative melanin gene cluster of *M. graminicola* (*Z. tritici*), but has not been functionally characterised (Lendenmann et al., 2014). Therefore, Ds37856 may have a similar function for melanin/elsinochrome biosynthesis in *D. septosporum* and may explain why it was upregulated at an earlier stage of infection than the other three genes.

The complete pathway for elsinochrome biosynthesis is not yet known. On the other hand, melanin biosynthesis is a well-studied subject. There are two major biosynthetic pathways for fungal melanin biosynthesis. The 1,8-dihydroxynaphthalene (DHN) melanin biosynthesis pathway is generally found in ascomycete fungi, the group to which *D. septosporum* belongs, whilst the L 3–4 dihydroxyphenylalanine (L-DOPA) melanin pathway is generally found in basidiomycetes (Bell and Wheeler, 1986). Both of these pathways are well characterised. The DHN-melanin biosynthesis requires a PKS, multicopper oxidase (also called laccase or p-diphenol oxidase in gene clusters), T3HN/T4HN reductases, and scytalone dehydratase (Langfelder et al., 2003, Eliahu et al., 2007). In contrast the DOPA pathway produces melanin from tyrosine without the use of
a PKS (Eisenman and Casadevall, 2012). Figure 3.29 shows the fungal DHN-melanin biosynthesis pathway in *C. heterostrophus*. Melanin gene clusters of *C. heterostrophus*, *A. fumigatus*, and *M. grisea* contain genes predicted to encode similar proteins such as reductases and dehydratases, but also have genes with different predicted functions (Langfelder et al., 2003; Eliahu et al., 2007). For example, *C. heterostrophus* and *M. grisea* melanin gene clusters have genes encoding zinc-finger transcription factors, however, the *A. fumigatus* melanin cluster has no gene with a similar function (Eliahu et al., 2007). The *BRN1* gene of *C. heterostrophus* is predicted to encode a 1,3,8-trihydroxynaphthalene (T3HN) reductase, an enzyme that converts T3HN to vermelone (Figure 3.29), and its closest *D. septosporum* homolog is *DsBrn1*, a gene in the putative *DsPks1* gene cluster (amino acid identity: 78.3%, E=2.7e-139). In addition, the third best BlastP hit of the TH4N reductase gene ChBRN2 is DsBrn1 (identity: 47%, E=2.2e-87), suggesting similar structures of the two reductases in melanin biosynthesis.

In order to determine the possibility of a fragmented *DsPks1* gene cluster, *D. septosporum* proteins with similar amino acid sequences to the melanin cluster gene products in *C. heterostrophus* and *A. fumigatus* were found using BlastP. Figure 3.30 shows candidate genes that may be involved in the biosynthesis of *DsPks1* SM, and their *in planta* expression levels. A key melanin biosynthetic gene that is not found in the *DsPks1* cluster, encoding scytalone dehydratase, is present in the *D. septosporum* genome on chromosome 3 (Ds70291). This single-copy gene is homologous to *C. heterostrophus* and *A. fumigatus* scytalone dehydratases (ChSCD1 and AfARP1), with 73.8% amino acid identity to ChSCD1 (E=2e-92). It is expressed at a significantly higher level at the late stage of infection, compared to earlier stages, similar to the *DsPks1* putative cluster genes (Figure 3.30).

Other melanin biosynthetic genes that are not found in the *DsPks1* gene cluster but are present in the *C. heterostrophus* and *A. fumigatus* melanin clusters are T4HN reductase and multicopper oxidase/laccase. Among all *D. septosporum* genes whose products have similar amino acid sequences to these, two pairs are physically in close proximity to each other. Ds69238 (multicopper oxidase) and Ds69231 (T4H reductase) are in chromosome 2 with 6 kb and 1 interfering gene
between them. Ds69238 is the best BlastP hit of multicopper oxidase AfABR1 in *D. septosporum* (45.5% identity, E=8e-100) and expressed at a significantly higher level at late stage of infection compared to early and mid stages. Ds69231 has 39% amino acid identity to T4HN reductase ChBRN2 (E=2.5e-7), but most highly expressed at mid stage and downregulated at late stage. Another candidate gene pair, Ds129162 (multicopper oxidase) and Ds87983 (T4HN reductase), is in chromosome 4 with 9 kb and 3 interfering genes between them. Ds129162 has 42.3% amino acid identity to AfABR1 (E=2.16e-40) and had a higher expression level at late stage than early and mid stages. Ds87983 was also expressed at a higher level at late stage compared to early and mid stages, and amino acid identity of Ds87983 to ChBRN2 is 35.5% (E=3.4e-12). However, overall *in planta* expression levels (RPMK values) of Ds129162 and Ds87983 were lower than that of the putative *DsPks1* gene cluster. Positions and BlastP results of the candidate fragmented cluster genes are summarized in Table 3.8.

![Figure 3.28. Melanin biosynthesis pathway in *C. heterostrophus*. The enzymes predicted to be involved in the biosynthesis are PKS18 (polyketide synthase), BRN2 (T4HN reductase), BRN1 (T3HN reductase), SCD1 (scytalone dehydratase), and p-diphenol oxidase are enzymes presumably involved in the indicated biosynthetic steps. Proteins with similar functions to the products of the genes in *DsPks1* putative gene cluster are marked with asterisk. Image reproduced from Eliahu et al. (2007).](image)
Figure 3.29. Expression of candidate genes in the putative fragmented DsPks1 gene cluster. Significant differences in the gene expression between early-mid and mid-late stages of infection were marked with "*" and "+", respectively (p<0.05). The differences between early and late stages were significant for all genes (p<0.05). The genes in the original gene cluster from Figure 3.28 are indicated with asterisk. Predicted functions of the candidate genes outside of the original gene cluster are shown in brackets as MCO: multicopper oxidase, T4HN: T4HN reductase, SCD: scytalone dehydratase. Candidate gene pairs that are in the same chromosome are closer together.

Table 3.8. Candidate genes involved in melanin biosynthesis outside of DsPks1 original gene cluster.

<table>
<thead>
<tr>
<th>JGI protein IDa</th>
<th>Scaffold: position (direction)b</th>
<th>Similar gene (function)c</th>
<th>E / identity %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ds69238</td>
<td>2: 661823-666821 (-)</td>
<td>AFABR1 (MCO)</td>
<td>8e-100/45.5</td>
</tr>
<tr>
<td>Ds69231</td>
<td>2: 656938-658277 (+)</td>
<td>ChBRN2 (T4HN)</td>
<td>2.5e-7/39</td>
</tr>
<tr>
<td>Ds129162</td>
<td>4: 2117920-2119714 (+)</td>
<td>AFABR1 (MCO)</td>
<td>2.2e-40/42.3</td>
</tr>
<tr>
<td>Ds87983</td>
<td>4: 2128494-2131693 (+)</td>
<td>ChBRN2 (T4HN)</td>
<td>3.4e-12/35.5</td>
</tr>
<tr>
<td>Ds70291</td>
<td>3: 1012112-1013484 (-)</td>
<td>ChSCD1 (SCD)</td>
<td>1.9e-92/73.8</td>
</tr>
</tbody>
</table>

b Position (nt) of predicted gene in scaffold according to JGI.
c Genes with similar predicted roles to the candidate DsPks1 cluster genes. Putative functions of the genes are shown in brackets as MCO: multicopper oxidase, T4HN: T4HN reductase and SCD: scytalone dehydratase.

In conclusion, the DsPks1 putative gene cluster is most likely responsible for the biosynthesis of a DHN-melanin, or possibly elsinochrome. If Pks1 SM is a DHN-melanin there may be genes encoding for scytalone dehydratase, T4HN reductase and multicopper oxidase involved in its biosynthesis that are located outside of the gene cluster. There is one candidate for scytalone dehydratase and at least two
candidates for each of T4HN reductase and multicopper oxidase encoding genes, some of which are grouped together on a separate chromosome, suggesting that DsPks1 may be a fragmented gene cluster.

In summary, phylogeny and gene cluster analyses suggest that DsPks1 may be involved in melanin biosynthesis. Genes potentially involved in melanin biosynthesis appear to be present outside of as well as in the DsPks1 putative gene cluster (Figure 3.31). In addition, DsPks1 has a correct gene model and is under strong negative selection.

**Figure 3.30. Candidate genes that may belong to a putative DsPks1 fragmented gene cluster for melanin biosynthesis.** Numbers inside arrows indicate JGI protein IDs. Pink: Secondary metabolite core gene (DsPks1). Purple: putative DsPks1 cluster genes. Small white arrows: genes without predicted SM functions. “Ch” numbers indicate the chromosomes in which the genes are located. Predicted functions of the candidate genes outside of the original gene cluster are shown in under the genes as MCO: multicopper oxidase, T4HN: T4HN reductase, SCD: scytalone dehydratase.
3.3.2. PKS2

The expression level of \textit{DsPks2} was higher than that of the dothistromin core gene, \textit{DsPksA}, at the late stage of plant infection (Section 3.1.2). Therefore, \textit{DsPks2} was a gene of main interest in this project so, in addition to the analyses done for all \textit{D. septosporum} SM core genes, phylogenetic analysis and gene knockout attempts were done with \textit{DsPks2}.

3.3.2.1. Phylogenetic analysis of DsPks2

Phylogenetic analysis of DsPks2 was done to find out if there are any putative orthologs that are functionally characterised. Initial phylogenetic analyses were done using KS domains, KS+AT domains, and full protein sequences as explained in Section 2.14.3. As a result of all three analyses, DsPks2, a protein belonging to a species of Dothideomycetes class was clustered with proteins of \textit{Aspergillus spp.}, members of the Eurotiomycetes class (Appendix 14). This suggested the possibility of horizontal gene transfer between \textit{D. septosporum} and \textit{Aspergillus spp.}. Next, phylogenetic analysis on DsPks2 was done using full protein sequences in order to determine true orthologs/paralogs instead of only determining proteins with similar KS and AT domains. In order to determine if there was a horizontal gene transfer, an extensive phylogenetic tree with a larger number of sequences compared to the initial tree, was prepared. The DsPks2 sequence was aligned with the best reciprocal FASTA hits from all Dothideomycetes, Eurotiomycetes and Sordariomycetes fungi present in the JGI database, and a tree was built using maximum likelihood (ML) method, as explained in Section 2.14.3. \textit{Fusarium} spp. (in Sordariomycetes) were included in the analysis since the SM produced by DsPks2 was predicted to be a fumonisin-like compound, a SM produced by \textit{Fusarium} spp. (Table 3.1; Rheeder et al., 2002).
Figure 3.31. Phylogenetic analysis of DsPks2 and its best FASTA hits. Red: Dothideomycetes, Blue: Eurotiomycetes, Green: Sordariomycetes. Numbers on branches indicate approximate likelihood ratio test (aLRT) scores. Close homologs of *D. septosporum* are shown with a red square, *D. septosporum* is indicated with an asterisk. Fusarium spp. are indicated with a plus sign. Details of the genome sequences used are presented in Appendix 23.
The phylogenetic tree in Figure 3.32 revealed that even though sequences belonging to the same fungal class are clustered at some regions of the tree as expected, many clades contain a mixture of sequences from different fungal classes. This is characteristic of a tree containing not only orthologous genes, but also many paralogs or distantly related homologs. Figure 3.33 shows the amino acid alignment of DsPks2 with closely related sequences shown in the red square of Figure 3.32. The closest homologs of DsPks2 are in Zopfia rhizophila (protein ID: 734285), Apiospora montagnei (protein ID: 166223), Lepidopterella palustris (protein ID: 437929), and Talaromyces acuelatus (protein ID: 470923), with amino acid identities to DsPks2 of 51.8% (E=0), 51.1% (E=0), 59.7% (E=0), and 51.2% (E=0) respectively. None of the close homologs of DsPks2 were functionally characterised, and DsPks2 was far from Fusarium spp. in the phylogenetic tree. Therefore this phylogenetic analysis did not support the hypothesis that Pks2 SM may be a compound similar to fumonisin produced by Fusarium spp. The alignment shows that there are only a few conserved regions between these five sequences, suggesting that these proteins are probably not orthologs.

In conclusion, the phylogenetic analysis of DsPks2 suggests that DsPks2 was probably not horizontally transferred and its close homologs are probably not true orthologs. This suggests that the metabolite produced by DsPks2 may be unique among Dothideomycetes, Eurotiomycetes, and Sordariomycetes.

### 3.3.2.2. Confirmation of gene model for DsPks2

The JGI gene model of DsPks2 was confirmed by aligning DsPks2 with its closest homologs from Section 3.3.2.1, and analysing the RNA-seq data.
The domain organization of DsPks2 (protein ID: 73814) and its alignment with closest homologs is presented in Figure 3.34. N-terminal sequences in the alignment were less conserved than the C-terminal sequences. The position of the first amino acid of DsPks2 was conserved with the first amino acids of the other three protein sequences in the alignment. There was an alignment gap in DsPks2 just after the start codon but this did not affect any of the known functional domains such as the KS domain. Therefore, the alignment of DsPks2 and its closest homologs support that the JGI gene model for DsPks2 is correct.

The predicted gene model for DsPks2 has 10 introns, and all of them were confirmed by RNA-seq data that was mapped to the JGI gene model of DsPks2 (Appendix 15). Therefore, alignments (Figure 3.34, Appendix 21e) and RNA-seq data (Appendix 15) support that the gene model for DsPks2 is correct.

3.3.2.3. Comparison of DsPks2 from 19 strains of D. septosporum

Evolutionary selection analysis was done by aligning the DsPks2 nucleotide sequences across the 19 D. septosporum strains and determining the selection pressure by dN/dS analysis and identifying codons that are under significantly strong positive selection pressure.

Figure 3.35 shows the alignment of DsPks2 across the 19 D. septosporum strains. The dN/dS ratio of DsPks2 is 0.21, suggesting that this gene is under negative selection. Therefore, the Pks2 SM may be an important metabolite for D. septosporum. One codon in the DH domain showed significantly positive selection (P>0.99) in DsPks2.
Figure 3.33. Domain structure of DsPks2 and its protein alignment with closest homologs. DsPks2 indicates the peptide’s domain organization with KS: keto-synthase, AT: acyltransferase, ER: enoylreductase, KR: ketoreductase, ACP: Acyl carrier protein domains. Numbers indicate amino acid positions of domain borders. Dotted lines show position of DsPks2 gap in the alignment. The total length is 2507 amino acids. The diagram was drawn to scale. The dark regions indicate the amino acid alignment with black (highly similar) to white (not similar) sequences with respect to the consensus sequence. Triangles indicate corresponding intron positions. The amino acid alignment below demonstrates Ds: D. septosporum, Zr: Z. rhizophila (protein ID: 734285), Am: A. montagnei (protein ID: 166223), Lp: L. palustris (protein ID: 437929). The coloured bar represents consensus identity with red (below 30%); green shades indicate higher levels of identity (up to 100%) with bar height proportional to % identity.
Figure 3.34. *DsPks2* nucleotide alignment of NZE10 with 18 additional strains. Red arrows show the domain positions with KS: keto-synthase, AT: acyltransferase, ER: enoylreductase, KR: ketoreductase, ACP: acyl carrier protein domains. The bold black line on the DH domain indicates the statistically significant positively selected codon (p<0.01). Yellow arrows indicate intron positions. The coloured bar represents consensus identity with brown indicating identities above 30% and below 100% and green indicating 100% identity. For each sequence, vertical bars on the top and bottom lines show differences in nucleotide and amino acid sequences compared to NZE10. A bar on the top line alone shows a synonymous mutation and a bar on both lines shows a non-synonymous mutation. Strains used in this analysis are listed in Appendix 5.
3.3.2.4. Gene cluster analysis

*DsPks2* gene cluster analysis was performed in order to find possible genes taking a role in the biosynthesis of the polyketide product. The *DsPks2* putative gene cluster is predicted to contain 10 genes in a 29 kb region of chromosome 8. The biggest distance between two genes within the putative gene cluster is 2.5 kb, containing 1 gene with no SM-related function. All genes except Ds36908 were expressed at a significantly higher level at the late stage of plant infection compared to the early stage, suggesting co-expression of the putative cluster genes.

All genes except Ds36908 have predicted functions similar to genes in other fungal PKS gene clusters. The putative *DsPks2* gene cluster contains two MFS transporters (Table 3.9). A gene encoding a MFS transporter is present in *Pks4* gene cluster responsible for bikaverin biosynthesis in *F. verticillioides* (Brown and Webber, 2008). Other genes within the putative gene cluster are probably responsible for the biosynthesis of the corresponding metabolite. A gene predicted to encode a short chain dehydrogenase/reductase is present in the fumonisnin gene cluster of *F. verticillioides* (Butchko et al., 2003). In the fumonisnin-like gene cluster of *A. niger*, *Fum10* gene is predicted to encode an AMP-dependent synthetase and ligase (Baker, 2006). Genes encoding for alpha-beta hydrolases are found in the SM gene cluster of the hybrid polyketide-nonribosomal peptide pseurotin A in *A. fumigatus* (Owens et al., 2014). In addition, in the polyketide lomaiviticin biosynthetic gene cluster of *Salinispora tropica*, there is a hypothetical protein that has as its closest homolog an alpha/beta hydrolase (Kersten et al., 2013). Acyttransferases are commonly found in fungal polyketide gene clusters, taking a role in either the biosynthesis of the metabolite or the release of the polyketide from a fungal megasynthase (Bonsch et al., 2016; Xie et al., 2009). A gene predicted to encode a beta-lactamase is found in a polyketide gene cluster of fungus *Usnea longissima*, responsible for anthraquinone biosynthesis (Wang et al., 2014).
Figure 3.35. Organization and expression of *DsPks2* and its neighboring genes in early, mid, and late stages of infection. Numbers inside arrows indicate JGI protein IDs. Pink: Secondary metabolite core gene (*DsPks2*). Purple: putative *DsPks2* cluster genes. Small white arrows: genes without predicted SM functions. Significant differences between early-mid and mid-late stages were marked with "**" and "+", respectively (p<0.05). The differences between early and late stages were also significant for all cluster genes except *Ds36908*.

Table 3.9. Genes within *DsPks2* putative gene cluster.

<table>
<thead>
<tr>
<th>Protein IDa</th>
<th>Predicted proteinb</th>
<th>Best hit organism/JGI ID</th>
<th>E/ identity %</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Ds73810</td>
<td>Major facilitator transporter</td>
<td>Apiospora montagnei/149305</td>
<td>1.7e-162/52.6</td>
</tr>
<tr>
<td>*Ds55316</td>
<td>Short-chain dehydrogenase/reductase</td>
<td>Lepidopterella palustris/365300</td>
<td>3.2e-98/55</td>
</tr>
<tr>
<td>*Ds176708</td>
<td>Major facilitator transporter</td>
<td>Lepidopterella palustris/420042</td>
<td>0/59.4</td>
</tr>
<tr>
<td>*Ds64818</td>
<td>AMP-dependent synthetase and ligase</td>
<td>Cladosporium fulvum/192541</td>
<td>0/57.5</td>
</tr>
<tr>
<td>*Ds73814/Pks2</td>
<td>Polyketide synthase</td>
<td>Lepidopterella palustris/437929</td>
<td>0/59.2</td>
</tr>
<tr>
<td>*Ds90348</td>
<td>Alpha/beta hydrolase</td>
<td>Lepidopterella palustris/409407</td>
<td>5.9e-114/60.1</td>
</tr>
<tr>
<td>*Ds36907</td>
<td>Acyltransferase</td>
<td>Decorospora gaudfroyi/617224</td>
<td>1.6e-99/47.6</td>
</tr>
<tr>
<td>*Ds36908</td>
<td>Squalene/phytoene synthase</td>
<td>Zymoseptoria pseudotritici/796149</td>
<td>1.4e-72/64.3</td>
</tr>
<tr>
<td>*Ds73815</td>
<td>AMP-dependent synthetase and ligase</td>
<td>Zasmidium cellare/66809</td>
<td>4.8e-131/51.5</td>
</tr>
<tr>
<td>*Ds73816</td>
<td>Beta-lactamase</td>
<td>Lepidopterella palustris/309548</td>
<td>2.3e-156/47.7</td>
</tr>
</tbody>
</table>

b Predicted proteins according to antiSMASH and InterProScan
Asterisk indicates genes with similar predicted functions present in other fungal PKS gene clusters.
All BlastP hits were confirmed true except *Ds73810, Ds64818, Ds36908 and Ds73815* with reciprocal blast analyses.
A gene predicted to encode squalene/phytoene synthase is also present in the \textit{DsPks2} putative gene cluster. Genes encoding for squalene/phytoene synthases are commonly found SM gene clusters, but instead of polyketides, they are mostly found in terpene gene clusters (Song et al., 2014; Jeandet et al., 2013).

According to NaPDoS predictions (Section 2.14.2), \textit{DsPks2} may be responsible for the synthesis of a fumonisin-like compound (Table 3.1). Biosynthetic gene clusters for fumonisin produced by \textit{Fusarium} spp. and fumonisin-like compounds produced by \textit{A. niger} are well characterised (Alexander et al., 2009, Baker et al., 2006). The predicted roles of the genes in both of these gene clusters have variations between each other and with the \textit{DsPks2} putative gene cluster. However, there are three genes common in both \textit{Fusarium} spp. and \textit{A. niger} fumonisin gene clusters that appear to be missing in the \textit{DsPks2} putative gene cluster. These genes are cytochrome p450 monooxygenase, dioxygenase, and alcohol dehydrogenase. Therefore, if the \textit{DsPks2} SM is fumonisin-like, it is possible that genes with these three functions that are located outside of the gene cluster may be involved in its biosynthesis.

### 3.3.2.5. Gene knockout attempts

Exhaustive attempts were made to obtain a gene knockout to functionally characterise the highly expressed SM gene \textit{DsPks2}. For that, a gene knockout construct was prepared using the GATEWAY system, and protoplast transformation was performed. In four separate transformation attempts, a total of 219 transformed colonies were screened and only one gene KO candidate was found (~0.45% efficiency of integration into correct site). Unfortunately, this candidate was found to contain both wild type and \textit{ΔDsPks2} gDNAs, and even after three steps of purification that candidate did not give a pure strain with the \textit{ΔDsPks2} knockout genotype. It is possible that a functional copy of \textit{DsPks2} is essential for viability, so the knockout mutant was forced to remain as a heterokaryon with a wild-type strain but eventually this aspect of the project was abandoned. PCR and Southern blot results are presented in Appendix 17.
To summarize, *DsPks2* is not horizontally transferred, has a correct JGI gene model, and in a putative cluster of 10 genes. Gene KO attempts for *DsPks2* for functional characterization were unsuccessful and eventually abandoned.
3.3.3. PKS3

3.3.3.1. Confirmation of gene model for DsPks3

Confirmation of the JGI gene model for DsPks3 was done by aligning DsPks3 protein sequence with its closest homologs and analysing the RNA-seq reads. Figure 3.37 shows the protein alignment. Amino acid identities of DsPks3 with closest reciprocal BlastP hits were 59.5% (E=0) for the PKS in Melanconium sp. (protein ID: 89670), 57.9% (E=0) for that of M. grisea (protein ID: 116015), and 47.8% (E=0) for Talaromyces aculeatus PKS (protein ID: 460631). The alignment has an overall low consensus, especially in the region between DH and ER domains. Positions of 7 of the 10 introns of DsPks3 are represented as gaps in the amino acid alignment, suggesting possible misannotation of those introns. RNA-seq analysis (Appendix 18) showed that all introns except the 7th intron are misannotated, shown by the in culture expression of the gene at the intron positions. Intron 7 was correctly annotated, but introns 8 and 9 are shorter than was shown in the gene model. Strikingly, there are many internal stop codons in all three reading frames, as well as frameshifts, if all introns except 7, 8, and 9 are removed and counted as exon sequence (Appendix 18), suggesting that DsPks3 is a pseudogene. In addition, alignment of DsPks3 with the closest BlastP hits after removing the first 7 introns removed the gaps in the alignment; however, this also decreased the amino acid identity of the alignment and generated new gaps (Figure 3.37b). This raised possibility of introns that weren’t annotated in the gene model and decreased amino acid identity caused by the changing reading frame.
Figure 3.36. Domain structure of DsPks3 and its protein alignment with closest homologs. DsPks3 indicates the peptide’s domain organization with KS: keto-synthase, AT: acyltransferase, DH: dehydratase, ER: enoylreductase, KR: ketoreductase domains. Numbers indicate amino acid positions of domain borders in addition to start and end of the DsPks3 peptide sequence. a) The amino acid alignment based on the JGI gene model. Triangles indicate corresponding intron positions. Dotted lines show positions of the DsPks3 gaps in the alignment. Total length is 2127 amino acids. The diagram was drawn to scale. The dark regions indicate the amino acid alignment with black (highly similar) to white (not similar) sequences with respect to the consensus sequence. Ds: D. septosporum Pks3, Ms: Melanconium sp. (protein ID: 89670), Mg: M. grisea (protein ID: 116015), Ta: T. aculeatus (protein ID: 460631). The coloured bar represents consensus identity with red (below 30%); green shades indicate higher levels of identity (up to 100%) with bar height proportional to % identity. b) The amino acid alignment after removing the all of the introns except introns 7,8 and 9.
In conclusion, alignments and RNA-seq reads do not support the current JGI gene model for *DsPks3* but suggest that all introns except intron 7 are misannotated. Among the misannotated introns, introns 8 and 9 are probably shorter than they were shown in the gene model and the other misannotated introns are probably parts of the coding region. The results also suggest that *DsPks3* has undergone pseudogenization by the number of internal stop codons in the coding sequence after removal of the misannotated introns.

### 3.3.3.2. Comparison of *DsPks3* from 19 strains of *D. septosporum*

Figure 3.38 shows the alignment of *DsPks3* nucleotide sequences across *D. septosporum* NZE10 and 18 other strains. This alignment was based on the JGI gene model and shows that two strains RUS1 and DEN1 have internal stop codons due to point mutations. In addition, *DsPks3* sequences were truncated from the 5’ ends in two strains, COLS and GUA2. No sequences similar to the 5’ end of *DsPks3* were found in the 5’ flanking regions of the truncated regions of these two sequences. The additional evidence for pseudogenisation events amongst 4 of the strains studied here, suggested *DsPks3* is probably not a functionally important gene for *D. septosporum*, therefore no further analyses were done.
Figure 3.37. *DsPks3* nucleotide alignment of NZE10 with 18 additional strains. Red arrows show the domain positions with KS: keto-synthase, AT: acyltransferase, DH: dehydratase, ER: enoylreductase, KR: ketoreductase. Yellow arrows show intron positions. Internal stop codons in RUS1 and DEN1 are shown with bold blue lines. The coloured bar represents consensus identity with red indicating identities below 30%, green shades indicating above 30% and below 100% and green indicating 100% identity. For each sequence, vertical bars on the top and bottom lines show differences in nucleotide and amino acid sequences compared to NZE10. Black vertical regions in the alignments show "N" series. A bar on the top line alone shows a synonymous mutation and a bar on both lines shows a non-synonymous mutation except the "N" series where mutation analysis was not possible. Strains used in this analysis are listed in Appendix 5.
3.4. Hybrid polyketide-nonribosomal peptide synthetases

*D. septosporum* has two predicted hybrid polyketide-nonribosomal peptide synthetase (HPS) core genes (Table 3.1). However, expression levels of both these genes were very low both in culture and *in planta* (Section 3.1.2). Therefore, these two genes were not of key interest and only gene model confirmation, evolutionary selection pressure and gene cluster analyses were done on these genes.

3.4.1. HPS1

3.4.1.1. Confirmation of gene model for *DsHps1*

The *DsHps1* JGI gene model was confirmed by aligning *DsHps1* with best reciprocal BlastP hits and analysing RNA-seq data that was mapped on the gene model. The top three BlastP hits were to proteins of *Aspergillus terreus* and *Aspergillus sydowii* (Eurotiomycetes) and *Ophiostoma piceae* (Sordariomycetes). The *DsHps1* alignment (Figure 3.39) showed low overall consensus, with amino acid identities to these proteins of only 47.6% (*A. terreus*; E=0), 47.1% (*A. sydowii*; E=0), and 46.1% (*O. piceae*; E=0). The best BlastP hit of *DsHps1* within the Dothideomycete fungi was a protein from *Sporormia fimetaria* (protein ID: 458206) with 41.9% identity (E=6.8e-179). In the alignment *DsHps1* and its homologue in *A. terreus* had no big gaps or insertions, however C-termini of the proteins in *A. sydowii* and *O. piceae* were not aligned with *DsHps1*. Domain analyses on the *A. sydowii* and *O. piceae* protein sequences showed that they are PKS instead of HPS. Therefore, the differences between the sequences are probably not because of any problems with the *DsHps1* gene model. The *DsHps1* predicted gene model has no introns and the continuous RNA-seq reads from start codon to stop codon supported the lack of introns (Appendix 19). Therefore, alignments and RNA-seq reads supported the JGI gene model for *DsHps1*. 
Figure 3.38. Domain structure of DsHps1 and its protein alignment with closest homologs. DsHps1 indicates the peptide’s domain organization with KS: keto-synthase, AT: acyltransferase, DH: dehydratase, CM: C-methyltransferase, KR: ketoreductase, C: condensation, A: adenylation, T: thiolation domains. Numbers indicate amino acid positions of domain borders and the end of the peptide (positions of borders that are too close to another border are not shown). The total length is 3975 amino acids. The diagram was drawn to scale. The dark regions indicate the amino acid alignment with black (highly similar) to white (not similar) sequences with respect to the consensus sequence. The amino acid alignment shows Ds: D. septosporum Hps1, At: A. terreus (protein ID: 325), As: A. sydowii (protein ID: 25738), Op: O. piceae (protein ID: 4838). The coloured bar represents consensus identity with red (below 30%); green shades indicate higher levels of identity (up to 100%) with bar height proportional to % identity.
3.4.1.2. Comparison of DsHps1 from 19 strains of D. septosporum

Determination of the evolutionary selection pressure on DsHps1 was done by aligning the DsHps1 sequences across the 19 D. septosporum strains and determining the dN/dS ratio. The DsHps1 alignment is presented in Figure 3.40. The dN/dS ratio of DsHps1 is 0.39, suggesting negative selection on this gene. Since DsHps1 was not a gene of key interest due to its low expression levels both in culture and in planta (Section 3.1.2), sites under positive selection were not determined.
Figure 3.39. *DsHps1* nucleotide alignment of NZE10 with 18 additional strains. Red arrows show the domain positions with KS: keto-synthase, AT: acyltransferase, DH: dehydratase, CM: C-methyltransferase, KR: ketoreductase, C: condensation, A: adenylation, T: thiolation. The coloured bar represents consensus identity with brown indicating identities above 30% and below 100% and green indicating 100% identity. For each sequence, vertical bars on the top and bottom lines show differences in nucleotide and amino acid sequences compared to NZE10. A bar on the top line alone shows a synonymous mutation and a bar on both lines shows a non-synonymous mutation. Strains used in this analysis are listed in Appendix 5.
3.4.1.3. Gene cluster analysis

DsHps1 and DsDma1 genes are physically next to each other, therefore the associated putative gene cluster may contain both of these genes. Since DsHps1 and DsDma1 genes both have very low expression levels both in culture and in planta (Section 3.1.2), it was not possible to estimate the gene cluster by co-expression analysis. The putative DsHps1 - DsDma1 gene cluster (Figure 3.41) is composed of 9 genes in a 44 kb region of chromosome 11. Between these 9 genes, there are 7 more genes with no predicted SM-related functions. The biggest distance between two genes is 6.9 kb, containing 3 interfering genes (Figure 3.41). No genes with predicted SM-related functions were found among the 20 genes downstream of DsHps1.

![Figure 3.40. DsHps1 - DsDma1 putative gene cluster. Numbers inside arrows indicate JGI protein IDs. Pink: Secondary metabolite core genes (DsDma1 and DsHps1). Purple: putative DsDma1-DsHps1 cluster genes. Small white arrows: genes without predicted SM functions.](image)

**Table 3.10.** Genes within DsHps1 - DsDma1 putative gene cluster.

<table>
<thead>
<tr>
<th>JGI Protein IDa</th>
<th>Predicted proteinb</th>
<th>Best hit organism/JGI ID</th>
<th>E / identity %</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Ds75212</td>
<td>Flavin-containing monooxygenase</td>
<td>Cladosporium fulvum/191534</td>
<td>0/92.1</td>
</tr>
<tr>
<td>*Ds47966/Mox1</td>
<td>Flavin-containing monooxygenase</td>
<td>Cladosporium fulvum/191527</td>
<td>0/83.5</td>
</tr>
<tr>
<td>*Ds111675</td>
<td>Cytochrome P450</td>
<td>Cladosporium fulvum/188144</td>
<td>4.9e-50/73.8</td>
</tr>
<tr>
<td>*Ds28620</td>
<td>Cytochrome P450</td>
<td>Cochliobolus miyabeanus/10327</td>
<td>5.2e-145/43.8</td>
</tr>
<tr>
<td>*Ds139328</td>
<td>SAM-dependent methyltransferase</td>
<td>Patellaria atrata/1013025</td>
<td>4.2e-72/33.7</td>
</tr>
<tr>
<td>*Ds75219</td>
<td>Cytochrome P450</td>
<td>Thozetella sp./746229</td>
<td>5.9e-80/45.9</td>
</tr>
<tr>
<td>Ds28624</td>
<td>Crotonyl-CoA reductase</td>
<td>Zymoseptoria pseudotritici/799177</td>
<td>8e-139/61.7</td>
</tr>
<tr>
<td>*Ds28625/Dma1</td>
<td>DMA</td>
<td>Colletotrichum somersetensis/470342</td>
<td>3.4e-68/45.2</td>
</tr>
<tr>
<td>*Ds180045/Hps1</td>
<td>Hybrid PKS-NRPS</td>
<td>Aspergillus terreus/325</td>
<td>0/47.6</td>
</tr>
</tbody>
</table>


Asterisk indicates that genes with similar predicted functions are present in other fungal HPS gene clusters.

b Predicted proteins according to antiSMASH and InterProScan analyses.

All BlastP hits were confirmed true except Ds111675 and Ds28624 with reciprocal blast analyses.

All genes in the DsHps1 - DsDma1 putative gene cluster except Ds28624 have predicted functions common among other fungal HPS gene clusters. The putative gene cluster for DsHps1 - DsDma1 contains two flavin-containing monooxygenases.
A gene encoding flavin-containing monooxygenase, *cheE* was verified to be involved in hybrid polyketide-nonribosomal peptide chaetoblobosin biosynthesis in *Penicillium expansum* by RNA silencing (Schümann and Hertweck, 2007). In addition, the *DsHps1 - DsDma1* putative gene cluster has three cytochrome p450 proteins. Cytochrome p450 proteins are present in a variety of hybrid PKS-NRPS gene clusters, such as that for tenellin in *Beauveria bassiana*, fusarin C in *F. moniliforme*, cytochalasin E and K in *A. clavatus*, fusaridione and equisetin in *F. heterosporum* (reviewed by Fisch, 2013). In the phomopsin hybrid PKS-NRPS gene cluster in *Phomopsis leptostromiformis*, the *PhomM* gene is predicted to encode a SAM-dependent methyltransferase (Ding et al., 2016), and Ds139328 is predicted to encode a protein with similar function in the *DsHps1 - DsDma1* cluster.

Ds28624 is predicted to encode a crotonyl-CoA reductase. Whilst crotonyl-CoA reductase genes are common in bacterial hybrid PKS-NRPS gene clusters, no similar examples could be found in fungal SM gene clusters (Weber et al., 2008; Seipke and Hutchings, 2013).

The *DsHps1* SM was predicted to be a compactin/bacitracin type according to NaPDoS analysis (Table 3.1), although this prediction did not take the *DsDma1* gene into account. Compactin is a polyketide produced by *Penicillium spp.* (Abe et al., 2002), whilst bacitracin is an antibiotic produced by bacteria (Konz et al., 1997). The compactin gene cluster in *Penicillium citrinum* has 4 genes with predicted functions such as transesterase that are not found in the *DsHps1 - DsDma1* putative gene cluster (Abe et al., 2002). In addition, compactin biosynthesis does not require an NRPS or DMA activity. Therefore, it is unlikely that the SM produced by *DsHps1 – DsDma1* gene cluster is a compactin-like compound.

Another possibility for the type of SM produced by the *DsHps1 – DsDma1* gene cluster is cyclopiazonic acid (CPA). CPA is a specific inhibitor of calcium-dependent ATPase in sarcoplasmic reticulum that results in altered cellular Ca$^{2+}$ levels in the cytoplasm (Seidler et al., 1989). Also, a very recent finding has shown that CPA is a virulence factor of *Aspergillus flavus* in maize ear rot disease (Chalivendra et al.,
2017). In *A. flavus* the CPA biosynthesis gene cluster involves both *Hps* and *Dma* genes (Chang and Ehrlich, 2011). Interestingly, the gene cluster also contains a FAD-dependent monooxygenase and genes with similar predicted functions are also present in the *DsHps1*–*DsDma1* putative gene cluster (Table 3.10). *A. flavus* *Hps* (*pks-nrps*), *Dma* (*dmat*), and FAD-dependent monooxygenase (*maoA*) gene knockout mutants each lost the ability to produce CPA, and therefore these genes were functionally characterised. There is also a gene predicted to encode a cytochrome p450 upstream of the *A. flavus* CPA gene cluster, but this gene was not functionally characterised (Chang and Ehrlich, 2011). Amino acid identity of *DsHps1* to *AfPks-Nrps* is 43.6% (E=0) and identity of *DsDma1* to *AfDmat* is 38.6% (E=2.6e-6). These suggest that the *DsHps1*–*DsDma1* putative gene cluster may be responsible for biosynthesis of a CPA-like SM.

To sum up, the JGI gene model of *DsHps1* appear to be correct and there is negative selection pressure on the *DsHps1* gene. The predicted gene cluster of *DsHps1*–*DsDma1* has 9 genes and they may be involved in biosynthesis of a cyclopiazonic acid-like SM.

### 3.4.2. HPS2

#### 3.4.2.1. Confirmation of gene model for *DsHps2*

The JGI gene model for *DsHps2* was confirmed by aligning *DsHps2* with best reciprocal BlastP hits and analysing RNA-seq reads that were mapped on the gene model. The top reciprocal BlastP hits of *DsHps2* were in *Khuskia oryzae* (a Sordariomycete), *Ophiobolus disseminans* and *Verruculina enalia* (Dothideomycetes) with amino acid identities to *DsHps2* of 56.92% (E=0), 56.49% (E=0) and 55.56% (E=0) respectively; this suggested metabolites with similar properties to the *DsHps2* SM may be present in both Dothideomycetes and Sordariomycetes. The alignment (Figure 3.42) shows C-termini of the *O. disseminans* and *V. enalia* HPS were truncated. However, supported by the continuous RNA-seq reads at the 3’ end of *DsHps2* (Appendix 20), there is no intron in the 3’ end of the gene after the region encoding the last T domain and this region is part of the coding sequence. RNA-seq reads also support all of the
predicted introns in the region encoding \textit{DsHps2}. Therefore, the \textit{DsHps2} JGI gene model is correct, supported by alignments and RNA-seq data.
Figure 3.41. Domain structure of DsHps2 and its protein alignment with closest homologs. DsHps2 indicates the peptide’s domain organization with KS: keto-synthase, AT: acyltransferase, DH: dehydratase, CM: C-methyltransferase, KR: ketoreductase, T: thiolation, C: condensation, A: adenylation domains. Numbers indicate amino acid positions of domain borders and the end of the peptide (positions of borders that are too close to another border are not shown). Dotted lines show position of the low consensus region in the alignment. The total length is 4029 amino acids. The diagram was drawn to scale. Triangles indicate corresponding intron positions. The dark regions indicate the amino acid alignment with black (highly similar) to white (not similar) sequences with respect to the consensus sequence. The amino acid alignment shows Ds: D. septosporum Hps2, Ko: K. oryzae (protein ID: 424330), Od: O. disseminans (protein ID: 360096), Ve: V. enalia (protein ID: 565861). The coloured bar represents consensus identity with red (below 30%); green shades indicate higher levels of identity (up to 100%) with bar height proportional to % identity.
3.4.2.1. Comparison of *DsHps2* from 19 strains of *D. septosporum*

Comparison of *DsHps2* sequences across the 19 *D. septosporum* strains were done by aligning the DsHps2 protein sequences from these strains and determining the evolutionary selection pressure using dN/dS ratio (Figure 3.43). The dN/dS ratio of *DsHps2* is 0.22, suggesting *DsHps2* is under negative selection. Therefore, the metabolite produced by DsHps2 may have an important role for *D. septosporum*. Positively selected codons were not determined in this analysis.
Figure 3.42 *DsHps2* nucleotide alignment of NZE10 with 18 additional strains. Red arrows show the domain positions with KS: keto-synthase, AT: acyltransferase, DH: dehydratase, CM: C-methyltransferase, KR: ketoreductase, T: thiolation, C: condensation, A: adenylation. Yellow arrows indicate intron positions. The coloured bar represents consensus identity with brown indicating identities above 30% and below 100% and green indicating 100% identity. For each sequence, vertical bars on the top and bottom lines show differences in nucleotide and amino acid sequences compared to NZE10. A bar on the top line alone shows a synonymous mutation and a bar on both lines shows a non-synonymous mutation. Strains used in this analysis are listed in Appendix 5.
### 3.4.2.2. Gene cluster analysis

DsHps2 gene cluster analysis was done to identify the genes that may be involved in the biosynthesis of the HPS product. The expression level of DsHps2 was very low both in planta and in culture (Section 3.1.2), and therefore co-expression of the putative cluster genes could not be analysed. The DsHps2 putative gene cluster contains 7 genes spanning a 40 kb region of chromosome 9. In addition to these 7 genes, there are 3 genes with no predicted function related to secondary metabolism. The biggest distance between two genes with SM-related functions is 4.1 kb, containing 2 interfering genes (Figure 3.44).

**Figure 3.43. DsHps2 putative gene cluster.** Numbers inside arrows indicate JGI protein IDs. Pink: Secondary metabolite core gene (Hps2). Purple: putative DsHps2 cluster genes. Small white arrows: genes without predicted SM functions.

### Table 3.11 Genes within DsHps2 putative gene cluster.

<table>
<thead>
<tr>
<th>JGI Protein IDa</th>
<th>Predicted proteinb</th>
<th>Best hit organism/JGI ID</th>
<th>E / identity %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ds74346</td>
<td>Carbamoyl phosphate synthase</td>
<td><em>Cladosporium fulvum</em> /187961</td>
<td>0/81.8</td>
</tr>
<tr>
<td>+Ds157675</td>
<td>Major facilitator transporter</td>
<td><em>Cladosporium fulvum</em> /187962</td>
<td>0/80.7</td>
</tr>
<tr>
<td>+Ds157678/Hps2</td>
<td>Hybrid PKS-NRPS</td>
<td><em>Khusia oryzae</em> /424330</td>
<td>0/56.9</td>
</tr>
<tr>
<td>+Ds82322</td>
<td>Dehydrogenase / Reductase</td>
<td><em>Verruculina enaia</em> /489491</td>
<td>7.7e-121/62.1</td>
</tr>
<tr>
<td>+Ds65421</td>
<td>ABC transporter related protein</td>
<td><em>Rhizoclosmatium globosum</em> /850377</td>
<td>0/37.5</td>
</tr>
<tr>
<td>+Ds37533</td>
<td>Cytochrome P450</td>
<td><em>Dothidotithia</em> /303483</td>
<td>2.5e-168/56.8</td>
</tr>
<tr>
<td>+Ds65424</td>
<td>Crotonyl-CoA reductase</td>
<td><em>Ophiobolus disseminans</em> /427746</td>
<td>2.8e-159/62.2</td>
</tr>
</tbody>
</table>


Asterisk indicates that genes with similar predicted functions are present in other fungal PKS gene clusters. Plus indicates that genes with similar predicted functions are present in other fungal NPS gene clusters.

b Predicted proteins according to antiSMASH and InterProScan

All BlastP hits were confirmed true except Ds65421 with reciprocal blast analyses.

All genes in the DsHps2 putative gene cluster except Ds74346 and Ds65424 are present in both fungal PKS and NRPS gene clusters, and genes with similar functions were found in the other predicted SM gene clusters of *D. septosporum*. The DsHps2 putative gene cluster contains five biosynthetic genes and two transporter genes, one MFS and one ABC transporter (Table 3.11). Genes predicted to encode MFS type transporters were also found in DsNps2 (Section 3.2.2.3) and DsPks2 (Section 3.3.2.4) putative gene clusters. The DsNps2 putative gene cluster
also contains a gene predicted to encode an ABC transporter. Functional analysis of an ABC gene in the fumonisin gene cluster of *Fusarium verticillioides*, *FUM19*, suggested that FUM19 might be responsible for the export of fumonisin from hyphae (Proctor et al., 2003). Another *DsHps2* cluster gene, *Ds82322*, is predicted to encode a short-chain dehydrogenase/reductase, and genes with similar functions were found in the putative gene clusters of *DsNps3*, *DsPks1* and *DsPks2* (Sections 3.2.3.4, 3.3.1.4, 3.3.2.4). *Ds37533* is predicted to encode a cytochrome p450, similar to one found in the *DsHps1 – DsDma1* putative gene cluster (Section 3.4.1.3).

The *DsHps2* putative gene cluster contains two genes that weren’t found in other fungal PKS, NRPS, or HPS gene clusters, but genes with similar predicted functions were observed in other *D. septosporum* SM putative gene clusters. A gene predicted to encode a carbamoyl phosphate synthase was found in the *DsNps1* putative gene cluster (Section 3.2.1.3), and one predicted to encode a crotonyl-CoA reductase was found in the *DsHps1 – DsDma1* putative gene cluster (Section 3.4.1.3).

NaPDoS analysis suggested that the *DsHps2* SM might be a compactin or tyrocidine (Table 3.1). However, as explained in Section 3.4.1.3, compactin biosynthesis requires additional biosynthesis genes but does not require an NRPS (Abe et al., 2002). On the other hand, tyrocidine is a SM produced by bacteria (Mootz and Marahiel, 1997). Therefore, *DsHps2* SM is unlikely to be either compactin or tyrocidine and further work is required to determine the nature of the *DsHps2* SM.

To sum up, the *DsHps2* gene model is correct; it is under negative selection, and in a putative cluster of 7 genes. However, no prediction could be made on the SM produced by the *DsHps2* gene cluster.
4. Conclusions and future work

Among the SMs produced by *D. septosporum*, only dothistromin had been studied extensively before this PhD project. The previous studies revealed that dothistromin is a virulence factor (Kabir et al., 2015) and its production involves a fragmented gene cluster (Chettri et al., 2013). Genome sequencing revealed that, in addition to the *DsPksA*, dothistromin biosynthesis core gene, there are 10 more SM core genes in the *D. septosporum* genome (de Wit et al., 2012). The aim of this project was to characterise the full set of SM core genes, and to determine if *D. septosporum* has other SM virulence factors and fragmented SM gene clusters. The key conclusions of this study are that *D. septosporum* may be producing at least one other SM virulence factor and may have fragmented gene clusters for the biosynthesis of some other SMs.

An overall analysis of the 11 SM core genes was done to get an overview of the capacity of *D. septosporum* for secondary metabolism. It was previously reported that among the 11 SM core genes of *D. septosporum*, only *DsPksA*, *DsPkS1* and *DsNps2* have *C. fulvum* orthologs, and *DsPks4* is truncated (de Wit et al., 2012). In this project, gene model confirmation revealed that the JGI gene model of *DsPks3* is not correct, and *DsPks3* is also a pseudogene. Strikingly, evolutionary selection pressure analyses on SM core genes revealed that all of the analysed PKS, NRPS, and HPS core genes of *D. septosporum* are under negative evolutionary selection, with dN/dS values ranging from 0.13 (*DsPks1*) to 0.39 (*DsHps1*). This suggests that all these genes may have important roles at some stages of the fungal life cycle.

In culture and *in planta* expression of *D. septosporum* SM core genes revealed that, in addition to *DsPksA*, three more genes, *DsPks1*, *DsPks2*, and *DsNps3* were expressed at high levels during infection. Therefore, these three genes were selected as genes of key interest as virulence factor candidates. Interestingly, phylogenetic analyses revealed that two of these genes, *DsPks2* and *DsNps3* have no orthologs in any of the species in three fungal classes whose genomes are currently available in the JGI database. On the other hand, *DsNps1*, *DsNps2*, *DsHps1* and *DsHps2* had very low expression levels both in culture and *in planta*. Low
expression levels for these *D. septosporum* SM core genes suggest that these genes do not have functions related with plant disease. However, it is possible that expression levels of these genes could be higher if different conditions for SM production were provided in culture, or the right timing, host, or part of the host were used. For example, some of the SM core genes may be expressed at high levels in spores or needles that have fallen on the forest floor. For example, a polyketide gleosporone is produced at high levels in the spores of plant pathogen *C. gloeosporioides* to prevent germination until the right conditions for host infection are present (Meyer et al., 1983). In *Aspergillus* spp. it was seen that competition with another microorganism significantly altered the metabolite production (Losada et al., 2009). It is possible that some of the *D. septosporum* SMs may be produced on the needles on the forest floor to gain competitive advantage against saprophytes.

Effects of different media conditions on *D. septosporum* metabolite production were also analysed. In these analyses, it was shown that use of NH₄ as sole nitrogen source had the most dramatic effects on the metabolite profiles by significantly changing the production of three metabolites in wild type and two metabolites in the Δ*LaeA* strain. In contrast, use of NO₃ as sole nitrogen source had no significant effects on the metabolite profiles obtained. Because one of the SM core genes (*DsNps2*) was predicted to produce a siderophore, the effects of iron limitation were analysed. However, there was no increase in the metabolites produced by cultures grown in media without iron compared to those with iron. This suggests that siderophore production in *D. septosporum* might not be affected by iron deficiency and it was not possible to predict which of the TLC bands might be the possible siderophore produced by *DsNps2* gene cluster. The effect of carbohydrate limitation was also tested, and the amounts of one metabolite in wild type and one in Δ*LaeA* were significantly decreased when the glucose concentration was decreased. This was opposite to expected, as carbon catabolite repression is one of the main regulators of fungal secondary metabolism. The main limitations of these analyses included the timing and methods of extraction. If the production of a particular SM is different at an earlier time point than the time of the analysis, it might be not be possible to detect a carbon catabolite repression
effect. Use of different solvent systems to those used in this analysis might enable other metabolites to be detected. Future work for better characterization of the media effects may include use of different time points, different metabolite extraction methods, and other variables such as different types of carbon sources used in the analysis.

Gene cluster analyses on the *D. septosporum* SM core genes revealed that some predicted genes appear to be located outside of the gene clusters for several SMs. For example, *DsPks1*, a gene most likely involved in melanin biosynthesis, is probably part of a fragmented gene cluster distributed across multiple chromosomes. Likewise *DsNps1* and *DsNps2* gene clusters are predicted to be fragmented. This supports the hypothesis that there are other fragmented gene clusters used in the SM biosynthesis of *D. septosporum*, in addition to the characterised dothistromin biosynthesis gene cluster. Gene clusters can be confirmed using two ways. First, the cluster genes can be knocked out and the metabolite production could be tracked by analysing the metabolites using a method such as TLC or high-pressure liquid chromatography (HPLC). However, none of the *D. septosporum* SMs except dothistromin have been chemically characterised from the metabolite extracts yet. Secondly, an easier method involves the use of master regulatory genes. It is known that expression of many of the SM core genes (*DsPks2, DsNps1, DsNps2, DsHps1* and *DsHps2*) is altered in a strain of *D. septosporum* that has a deletion of the *DsLaeA* global regulator gene (*ΔLaeA*) compared to the wild type strain (Chettri and Bradshaw, 2016). Based on studies with *Aspergillus* spp. (Bok and Keller, 2004; Amaiike and Keller, 2009), co-regulation means the expression of most clustered SM genes would also be altered in a *ΔLaeA* strain, so this can be used to help confirm gene clusters. The limitation with this method is that the gene clusters whose core genes are not regulated by the *DsLaeA*, such as *DsPks1*, can’t be confirmed using the *DsLaeA* deletion mutants.

Apart from the potential fragmented gene clusters, gene cluster analyses for the SM core genes also revealed that *DsHps1 - DsDma1* putative gene cluster might be responsible for production of cyclopiazonic acid, which was different from the NaPDoS prediction.
Comparison of 19 *D. septosporum* genomes revealed not only the evolutionary selection pressures as mentioned previously, but also the higher number of polymorphisms in four Colombia (COL1-2) and Guatemala (GUA1-2) strains compared to the other 15 *D. septosporum* strains. This concurs with the finding that the Guatemala (GUA1) strain was the most divergent strain in a dothistromin cluster gene sequence analysis (Bradshaw et al., 2013). Therefore, COL and GUA appear to be deeply divergent strains of *D. septosporum*.

Phenotypic analysis of *D. septosporum ΔDsNps3* strains showed that the Nps3 SM might be a virulence factor. This was supported by lower hyphal surface network on needle surfaces at an early stage of infection and lower fungal biomass in late stage lesions. The expression level of *DsNps3* at the early stage of infection was the highest of the SM core genes and, combined with the significant difference in the fungal surface network between ΔDsNps3 and wild type strains it is possible that the DsNps3 SM may be associated with early / biotrophic stage infection. In the hemibiotrophic plant pathogen *C. higginsianum*, two NRPS genes were also expressed highly at the biotrophic stage of infection but these genes have not been functionally characterised. In the current study, exhaustive efforts were made to determine the kind of NRP that might be produced by *DsNps3* using both bioinformatics and biochemical approaches. However, due to inconsistencies between different prediction softwares in A domain specificity determination and uncertainty of the cycle repetitions it wasn’t possible to determine the Nps3 SM bioinformatically. TLC analysis on ΔDsNps3 and wild type strains was done to attempt biochemical characterization of the Nps3 SM. Unfortunately, due to low level production of the Nps3 candidate SM and poor separation from other metabolites on the TLC plate, it wasn’t possible to biochemically characterise Nps3 SM. In order to confirm if Nps3 SM is a virulence factor, the pathogenicity assay of ΔDsNps3 will need to be repeated with a second knock out mutant and complemented strain. The nature of the Nps3 SM may be determined by over-expression of *DsNps3* gene cluster using other media conditions or regulatory genes, and then biochemical characterization might be done by TLC, HPLC, mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy analyses.
The key findings of this project for *D. septosorum* SM core genes are summarized in Table 4.1.

**Table 4.1. Key findings of this project for *D. septosorum***

<table>
<thead>
<tr>
<th>Gene</th>
<th>Expression</th>
<th>Metabolite</th>
<th>Unique</th>
<th>Fragmented cluster</th>
<th>Virulence</th>
<th>dN/dS</th>
</tr>
</thead>
<tbody>
<tr>
<td>DsPks1</td>
<td>High</td>
<td>Melanin</td>
<td>No</td>
<td>Yes</td>
<td>Unknown</td>
<td>0.13</td>
</tr>
<tr>
<td>DsPks2</td>
<td>High</td>
<td>Unknown</td>
<td>Yes</td>
<td>Unknown</td>
<td>Unknown</td>
<td>0.21</td>
</tr>
<tr>
<td>DsNps1</td>
<td>Low</td>
<td>Cyclosporin</td>
<td>Yes</td>
<td>Yes</td>
<td>Unknown</td>
<td>0.23</td>
</tr>
<tr>
<td>DsNps2</td>
<td>Low</td>
<td>Ferricrocin</td>
<td>No</td>
<td>Yes</td>
<td>Unknown</td>
<td>0.27</td>
</tr>
<tr>
<td>DsNps3</td>
<td>High</td>
<td>Unknown</td>
<td>Yes</td>
<td>Unknown</td>
<td>Yes</td>
<td>0.27</td>
</tr>
<tr>
<td>DsHps1</td>
<td>Low</td>
<td>Cyclopiazonic acid</td>
<td>No</td>
<td>No</td>
<td>Unknown</td>
<td>0.39</td>
</tr>
<tr>
<td>DsHps2</td>
<td>Low</td>
<td>Unknown</td>
<td>No</td>
<td>Unknown</td>
<td>Unknown</td>
<td>0.22</td>
</tr>
</tbody>
</table>

* a In planta expression at early, mid, or late stages of infection. High: above 10 RPMK. Low: below 10 RPMK.

* Predicted SMs produced by the SM core genes based on orthology and gene cluster analyses.

* Gene is unique among dothideomycetes.

* Appears to involve a fragmented gene cluster.

* According to the gene KO and phenotypic analyses, probably a virulence factor.

* dN/dS ratio indicating evolutionary selection. dN/dS values smaller than 1 indicate negative selection.

There are several future studies that can be based on the findings of this project. The most important of these are:

- Confirmation of the gene clusters, especially for *DsPks1*, *DsPks2*, and *DsNps3*.

- Confirmation of Nps3 SM as a virulence factor.

- Chemical characterization of the Nps3 SM.

Other future works based on this project include the following:

- Gene KO and functional characterization of the other SM core genes, in particular *DsPks1* and *DsPks2*.

- Chemical characterization of the other SMs.

Altogether, the results of this project showed that a hemibiotrophic fungus is perfectly capable of keeping a viable lifecycle with a very low number of SMs. Therefore, the current assumption about pathogens with a necrotrophic stage requiring more SMs than biotrophs (reviewed by Pusztahelyi et al., 2015) is not always correct. In addition, fragmented SM gene clusters may not be exceptional in plant pathogenic fungi, and there might be many more fragmented SM gene clusters waiting to be found. Finally, this project revealed that there are still many
unknown SMs, and learning more about them will greatly broaden our understanding of plant-pathogenic fungi interactions.
APPENDIX

Appendices that are on the CD

Appendix 10
Fluorescence microscopy images of WT and ΔDsNps3 \textit{D. septosporum} infected needles (Section 3.2.3.8).

Appendix 12
Binocular microscopy images of WT and ΔDsNps3 \textit{D. septosporum} infected lesions (Section 3.2.3.8).

Appendix 21
Alignments of predicted SM core proteins with best BlastP matches

Appendix 22
\textit{D. septosporum} SM core gene and amino acid sequences

Appendix 1 – Media

All media were prepared in ddH$_2$O and autoclaved at 121 °C for 15 minutes. Antibiotics were added after the media cooled down to 50°C.

\textbf{Luria broth (LB) (Bertani, 1951)}
5 g/L tryptone (Oxoid, UK), 2.5 g/L yeast extract (Gibco BRL), 5 g/L NaCl (Panreac, Spain)

\textbf{LB agar (Bertani, 1951)}
5 g/L tryptone (Oxoid), 2.5 g/L yeast extract (Gibco BRL), 5 g/L NaCl (Panreac), 1.5% (wt/vol) Agar (Neogen Acumedia, MI, USA)
DM (Bradshaw et al., 2000)
50 g/L bacto malt extract (BD, NJ, USA), 23 g/L nutrient broth (Oxoid)

DM agar (Bradshaw et al., 2000)
50 g/L bacto malt extract (BD), 23 g/L nutrient broth (Oxoid), 1.5% (wt/vol) Agar (Acumedia)

DSM (Bradshaw et al., 2000)
20 g/L bacto malt extract (BD), 5 g/L yeast extract (Gibco BRL), 1.5% (wt/vol) Agar (Acumedia)

PMMG (McDougal et al., 2011)

<table>
<thead>
<tr>
<th></th>
<th>Amount (g/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MgSO₄·7H₂O (Scharlau, Spain)</td>
<td>0.2</td>
</tr>
<tr>
<td>K₂HPO₄·3H₂O (VWR)</td>
<td>1</td>
</tr>
<tr>
<td>NH₄NO₃ (Sigma)</td>
<td>1</td>
</tr>
<tr>
<td>KCl (Sigma)</td>
<td>0.2</td>
</tr>
<tr>
<td>FeSO₄·7H₂O (Univar, IL, USA)</td>
<td>0.002</td>
</tr>
<tr>
<td>ZnSO₄·7H₂O (VWR)</td>
<td>0.002</td>
</tr>
<tr>
<td>Asparagine (Sigma)</td>
<td>2</td>
</tr>
<tr>
<td>Glucose (Univar)</td>
<td>3</td>
</tr>
<tr>
<td>Agar (Acumedia)</td>
<td>20</td>
</tr>
</tbody>
</table>

Pine needles were soaked overnight in ddH₂O [10%(wt/vol)] and this water was used for the media preparation. pH was adjusted to 6.2 prior to autoclave.

1.5% RG media
50 g/L malt extract (Oxoid), 23 g/L nutrient broth (Oxoid) and 273.8 g/L Sucrose (VWR), 1.5% (wt/vol) agar (Acumedia).

0.8% RG media
8 g/L bacteriological agar (Oxoid), 50 g/L malt extract (Oxoid), 23 g/L nutrient broth (Oxoid) and 273.8 g/L Sucrose (VWR), 0.8% (wt/vol) agar (Acumedia).
Appendix 2 – Buffers and Solutions

All buffers and solutions were prepared in ddH₂O unless stated otherwise and autoclaved at 121°C for 15 minutes and allowed to cool before use.

2% CTAB Buffer

2%(w/v) Hexadecyltrimethylammonium bromide (CTAB) (Sigma); 1% (w/v) Polyvinylpyrrolidone (PVP40) (Sigma); 1.4 M NaCl (Panreac); 20 mM EDTA (Sigma); 0.1 M Tris – HCl pH8.0. The solution was autoclaved at 121°C for 15 minutes and allowed to cool before use.

TE buffer

10 mM Tris (Carl Roth, Germany), 1 mM EDTA (Sigma), pH8.0.

10x TBE Buffer

108 g/L Tris (Carl Roth), 55 g/L Boric Acid (Univar), 7.44 g/L EDTA, pH8.2.

OM buffer

1.4 M MgSO₄ and 10 mM Na₂HPO₄ (VWR) was dissolved in 100 mL Milli-Q water. pH was adjusted to 5.8 with 100 mM NaH₂PO₄,2H₂O (VWR). Dissolving to a final volume of 300 mL, ddH₂O was added. The solution was autoclaved at 121°C for 15 minutes and allowed to cool before use.

ST buffer

0.6 M sorbitol (Sigma), 100 mM Tris (Carl Roth) pH 8.0. The solution was autoclaved at 121°C for 15 minutes and allowed to cool before use.

STC buffer

1 M sorbitol (Sigma), 50 mM Tris (Carl Roth) pH 8.0, 50 mM CaCl₂ (Merck). The solution was autoclaved at 121°C for 15 minutes and allowed to cool before use.
6x loading dye
0.25% (wt/vol) bromophenol blue (Sigma), and 40% sucrose was filter sterilized using 0.2 μm syringe filter.

Ethidium bromide (EtBr)
From 10mg/mL stock solution, 100 μL EtBr was dissolved per 1L water.

Southern blot solution 1 (Depurination)
0.25 M HCl

Southern blot solution 2 (Denaturation)
0.5 M NaOH/0.5M NaCl (Panreac)

Southern blot solution 3 (Neutralization)
0.5M Tris (Carl Roth) pH7.4, 2 M NaCl (Panreac)

20 x SSC buffer
3 M NaCl (Panreac), 0.3 mM Na₃C₆H₅O₇ (Merck)

Buffer I: 0.1 M Tris, 0.15M NaCl (Panreac) pH7.5 (Southern blot, Hybridization). The solution was autoclaved at 121°C for 15 minutes and allowed to cool before use.

Buffer II: 1g skim milk/100 mL Buffer I (Southern blot, Hybridization)

Buffer III: 0.1 M Tris, 0.1 M NaCl (Panreac) pH9.5 (Southern blot, Hybridization). The solution was autoclaved at 121°C for 15 minutes and allowed to cool before use.

Washing solution 1: (2 x SSC, 0.1% SDS (VWR)) (Southern blot, washing)
Washing solution 2: (0.5 x SSC, 0.1% SDS (VWR)) (Southern blot, washing)
### Appendix 3 – Primers

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Lab number</th>
<th>Primer Sequence (5’-3’)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPS3-5FP-attb2R</td>
<td>1565</td>
<td>GGGGACAGCTTTCTTGTAACAAAGTGGCTGAGTAATACCTCCCAGGCTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGGGACTGCTTTTTTGTACAAACTTGCCAAATGAGCTGAG</td>
</tr>
<tr>
<td>NPS3-5RP-attb1R</td>
<td>1566</td>
<td>GGGGACAGCTTTCTTGTAACAAACTTGCCAAATGAGCTGAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGGGACTGCTTTTTTGTACAAACTTGCCAAATGAGCTGAG</td>
</tr>
<tr>
<td>NPS3-3FP-attb4_domain</td>
<td>1567</td>
<td>GGGGACAGCTTTCTTGTAACAAACTTGCCAAATGAGCTGAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGGGACTGCTTTTTTGTACAAACTTGCCAAATGAGCTGAG</td>
</tr>
<tr>
<td>NPS3-3RP-attb3_domain</td>
<td>1568</td>
<td>GGGGACAGCTTTCTTGTAACAAACTTGCCAAATGAGCTGAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGGGACTGCTTTTTTGTACAAACTTGCCAAATGAGCTGAG</td>
</tr>
<tr>
<td>NPS3-FP-wtcheck</td>
<td>1587</td>
<td>CCAGAAATGTAACCTGCGCTT</td>
</tr>
<tr>
<td>D. septo NRPS 3 reverse</td>
<td>1099</td>
<td>CTGTCGGGACATTTGAGTC</td>
</tr>
<tr>
<td>NPS3-5FP-screen</td>
<td>1617</td>
<td>TCTCAGACGTGGCAGAGTC</td>
</tr>
<tr>
<td>OscHygR</td>
<td>1412</td>
<td>GCCGATGCAAAGTGGCGATAAAAAC</td>
</tr>
<tr>
<td>NPS3-3RP-domain screen</td>
<td>1618</td>
<td>TCACGAGAGACAGGAGGAGGACT</td>
</tr>
<tr>
<td>OscHygF</td>
<td>1411</td>
<td>AGAGCTTTGGTACGCGGAAATTTGCG</td>
</tr>
<tr>
<td>PKS2-73814-KO5F</td>
<td>1315</td>
<td>GGGGACAGCTTTCTTGTAACAAACTTGCGGAGG</td>
</tr>
<tr>
<td>PKS2-73814-KO5R</td>
<td>1316</td>
<td>GGGGACAGCTTTCTTGTAACAAACTTGCGGAGG</td>
</tr>
<tr>
<td>PKS2-73814-KO3F</td>
<td>1317</td>
<td>GGGGACAGCTTTCTTGTAACAAACTTGCGGAGG</td>
</tr>
<tr>
<td>PKS2-73814-KO3R</td>
<td>1318</td>
<td>GGGGACAGCTTTCTTGTAACAAACTTGCGGAGG</td>
</tr>
<tr>
<td>F-upst-Pks2-screen</td>
<td>1381</td>
<td>ACATACTGTTGGACGGCGCCAG</td>
</tr>
<tr>
<td>hph Fwd2</td>
<td>175</td>
<td>ATTTACATTGCGGCCGATTTGCGAG</td>
</tr>
<tr>
<td>GFP-Fwd</td>
<td>1365</td>
<td>GCATGGACAGGCTTGACAAAG</td>
</tr>
<tr>
<td>R-downst-Pks2-screen</td>
<td>1381</td>
<td>ACATACTGTTGGACGGCGCCAG</td>
</tr>
<tr>
<td>PKS2-ks-fp</td>
<td>1552</td>
<td>ACAGGAGAGCAATTGGCGCAG</td>
</tr>
<tr>
<td>D.septo pks2Reverse</td>
<td>1106</td>
<td>GATGTGATCAAGAATTGGCGCCAG</td>
</tr>
<tr>
<td>AflJ exF</td>
<td>1023</td>
<td>GACCTGTTGGCAGATTCGG</td>
</tr>
<tr>
<td>AflJ exR</td>
<td>1024</td>
<td>GCTGTAGTGACGGAATTCGA</td>
</tr>
<tr>
<td>CAD918</td>
<td>935</td>
<td>CACGAAAGAGATTGGAGGAGCCT</td>
</tr>
<tr>
<td>CAD1019</td>
<td>936</td>
<td>TTCAATACCCCCATCTGAGATCAAC</td>
</tr>
</tbody>
</table>

*Respective att sites of each primer are underlined.
Primers that were used for the same purpose were grouped together.
Appendix 4 – Gene model confirmation of DsNps1 (Section 3.2.1.1)

An amino acid alignment of DsNps1 with its best BlastP hits is presented in the CD (Appendix 21a). RNA-seq reads that were mapped onto the JGI gene model were used in an attempt to confirm the DsNps1 gene model (Section 3.2.1.1). But, as shown below, due to the very low expression levels of DsNps1 both in planta and in culture, there were insufficient reads to confirm the intron positions.

Figure A4. DsNps1 RNA-seq reads that were mapped onto the JGI gene model. Pink rows at top are DsNps1 RNA-seq reads at late, middle and early stages (in planta) and in culture (from Bradshaw et al, 2016). The numbers (eg 8.00) are RPMK values. The red row predicted coding sequence (from start codon to stop codon) based on the gene model from JGI. Green bars represent the ESTs originally used for gene annotation. PMMG and Rich media (DSM, V8, DM, PD) indicate the types of media used for growing fungal mycelia before RNA extraction. Clusters: indicate 454 EST sequences of RNA assembled with Newbler to create isotigs that were then aligned against the genomic assembly using the BLAT tool. ESTs: EST/cDNA sequences aligned against the genomic assembly using the BLAT tool. Bottom two green bars indicate the combination of the two “clusters” or “ESTs”.
Appendix 5 – *D. septosporum* strains used in evolutionary selection pressure analyses

Genome sequences of *D. septosporum* strains collected across the globe were used to determine evolutionary selection pressures on SM core genes.

**Table A5. Strains of *D. septosporum* for which genome sequences were available.**

<table>
<thead>
<tr>
<th>ID</th>
<th>Country</th>
<th>Region</th>
<th>Host</th>
<th>Collection year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1_ALP3</td>
<td>Germany</td>
<td>Bavarian Alps</td>
<td><em>P. mugo</em></td>
<td>1996</td>
</tr>
<tr>
<td>2_AUS4</td>
<td>Austria</td>
<td>Lower Austria</td>
<td><em>P. sylvestris</em></td>
<td>2004</td>
</tr>
<tr>
<td>3_BHU1</td>
<td>Bhutan</td>
<td>Yusipang, Thimphu dzongkhag</td>
<td><em>P. radiata</em></td>
<td>2005</td>
</tr>
<tr>
<td>4_CAN3</td>
<td>Canada</td>
<td>British Columbia</td>
<td><em>P. contorta v. latifolia</em></td>
<td>1996</td>
</tr>
<tr>
<td>5_CHI17</td>
<td>Chile</td>
<td>Unknown</td>
<td><em>P. radiata</em></td>
<td>2001</td>
</tr>
<tr>
<td>6_COLN</td>
<td>Colombia</td>
<td>Sonora EST-1-128, Northern zone</td>
<td><em>Pinus ELL x TAE</em></td>
<td>2011</td>
</tr>
<tr>
<td>7_COLS</td>
<td>Colombia</td>
<td>Don Miguel – Lote 54 – Southern Colombia</td>
<td><em>P. kesiya</em></td>
<td>2011</td>
</tr>
<tr>
<td>8_DEN1</td>
<td>Denmark</td>
<td>Copenhagen</td>
<td><em>P. aristata</em></td>
<td>2013</td>
</tr>
<tr>
<td>9_ECU13</td>
<td>Ecuador</td>
<td>Unknown</td>
<td><em>P. radiata</em></td>
<td>2001</td>
</tr>
<tr>
<td>10_GRE1</td>
<td>Greece</td>
<td>Northern Greece</td>
<td><em>P. burtia &amp; P. nigra</em></td>
<td>2012</td>
</tr>
<tr>
<td>11_GUA1</td>
<td>Guatemala</td>
<td>Unknown</td>
<td><em>P. tecumumani</em></td>
<td>1983</td>
</tr>
<tr>
<td>12_GUA2</td>
<td>Guatemala</td>
<td>Jalapa</td>
<td><em>P. oocarpa</em></td>
<td>2012</td>
</tr>
<tr>
<td>13_NZE2</td>
<td>New Zealand</td>
<td>Central North Island</td>
<td><em>P. radiata</em></td>
<td>1965</td>
</tr>
<tr>
<td>14_NZE8</td>
<td>New Zealand</td>
<td>Bay of Plenty</td>
<td><em>P. radiata</em></td>
<td>2004</td>
</tr>
<tr>
<td>15_RUS1</td>
<td>Russia</td>
<td>Unknown</td>
<td><em>P. sylvestris</em></td>
<td>2015</td>
</tr>
<tr>
<td>16_SAF4</td>
<td>South Africa</td>
<td>Limpopo</td>
<td><em>P. radiata needles</em></td>
<td>2002</td>
</tr>
<tr>
<td>17_SLI1</td>
<td>Slovakia</td>
<td>Unknown</td>
<td><em>P. sylvestris</em></td>
<td>2001</td>
</tr>
<tr>
<td>18_USA12</td>
<td>USA</td>
<td>Oregon</td>
<td><em>P. ponderosa</em></td>
<td>1983</td>
</tr>
</tbody>
</table>
Appendix 6 – Gene model confirmation of DsNps2 (Section 3.2.2.1)

An amino acid alignment of DsNps2 with its best BlastP hits is presented in the CD (Appendix 21b). RNA-seq reads that were mapped onto the JGI gene model were used in an attempt to confirm the DsNps2 gene model (Section 3.2.2.1). The intron at the 3’ end of DsNps2 could be confirmed using RNA-seq data. However, due to low number of reads at 5’ end of DsNps2, the intron at 5’ end of the gene could not be confirmed with RNA-seq analysis.

Figure A6. DsNps2 RNA-seq reads that were mapped onto the JGI gene model. Pink rows at top are DsNps2 RNA-seq reads at late, middle and early stages (in planta) and in culture (from Bradshaw et al, 2016). The numbers (e.g. 11.00) are RPMK values. The bold red row is predicted coding sequence (from start codon to stop codon) based on the gene model from JGI. Thin red row indicates untranslated regions (UTRs). Green bars represent the ESTs originally used for gene annotation. PMMG and Rich media (DSM, V8, DM, PD) indicate the types of media used for growing fungal mycelia before RNA extraction. Clusters: indicate 454 EST sequences of RNA assembled with Newbler to create isotigs that were then aligned against the genomic assembly using the BLAT tool. ESTs: EST/cDNA sequences aligned against the genomic assembly using the BLAT tool. The bottom two green bars indicate the combination of the two “clusters” or “ESTs”.

167
Appendix 7 – Gene model confirmation of DsNps3 (Section 3.2.3.2)

An amino acid alignment of DsNps3 with its best BlastP hits is presented in the CD (Appendix 21c). RNA-seq reads that were mapped onto the JGI gene model were used to confirm the DsNps3 gene model (Section 3.2.3.2). The DsNps3 gene model was confirmed by continuous RNA-seq reads from start codon to stop codon.

Figure A7. DsNps3 RNA-seq reads that were mapped onto the JGI gene model. Pink rows at top are DsNps3 RNA-seq reads at late, middle and early stages (in planta) and in culture (from Bradshaw et al, 2016). The numbers (eg 24.50) are RPMK values. The bold red row is predicted coding sequence (from start codon to stop codon) based on the gene model from JGI. Thin red rows indicate untranslated regions (UTRs). Green bars represent the ESTs originally used for gene annotation. PMMG and Rich media (DSM, V8, DM, PD) indicate the types of media used for growing fungal mycelia before RNA extraction. Clusters: indicate 454 EST sequences of RNA assembled with Newbler to create isotigs that were then aligned against the genomic assembly using the BLAT tool. ESTs: EST/cDNA sequences aligned against the genomic assembly using the BLAT tool. The bottom two green bars indicate the combination of the two “clusters” or “ESTs”.
Appendix 8 – *DsNps3*-pOSCAR gene knockout construct (Section 3.2.3.5)

The gene knock out of *DsNps3* was done using the plasmid construct shown below. This construct was prepared as in Section 2.7.2 and used to generate \(\Delta DsNps3\) mutants using protoplast transformation as in Section 2.9.

**Figure A8.** *DsNps3*-pOSCAR KO construct. Hph-TrpC shows the hygromycin resistance marker. *att* sites at the borders of flanking regions are shown. Sizes of the whole plasmid, flanking regions, and *hph-TrpC* are shown in brackets in bp. Primer pairs used in PCR amplification of 5' and 3' flanking sites were NPS3-5FP-attb2R/NPS3-5RP-attb1R and NPS3-3FP-attb4_domain/NPS3-3RP-attb3_domain, respectively. Primers NPS3-5FP-attb2R/NPS3-3RP-attb3_domain were used to confirm plasmid generation. Plasmid was generated based on Paz et al. (2011).
Appendix 9. Southern hybridization confirmation of ΔDsNps3

Genomic DNAs of wild type and ΔDsNps3 candidates were digested with EcoRV and, using a probe that could bind to both wild type and ΔDsNps3, the absence of DsNps3 was confirmed using Southern hybridization. Figure 3.17 shows the restriction enzyme recognition sites and probe binding positions on DsNps3 in wild type and ΔDsNps3.

![Southern hybridization confirmation of ΔDsNps3](image)

Figure A9. Southern hybridization confirmation of ΔDsNps3. EcoRV-digested wild type, ΔDsNps3 KO1, and ΔDsNps3 KO2 gDNAs (left) and Southern hybridization of these gDNAs (right) using the probe in Figure 3.17. Expected fragment sizes were 4559 bp for wild type and 8786 bp ΔDsNps3.

Southern hybridization analysis presented in Figure A19 showed that KO2 appears to have single copy integration with no wild type gDNA. There was no conclusive result present for KO1 because of the low gDNA concentration from the start. High background was because of the 3-hour exposure time that was required in order to detect the band belonging to KO2. No band belonging to wild type was observed for KO1 and KO2. Also, the band belonging to wild type was stronger than the other two. That was most likely due to inefficient gDNA extraction from KO1 and KO2 by use of plant gDNA extraction kits with fungal samples (Section 2.3.1). Wild type D. septosporum gDNA was present in the lab stock.
Appendix 11 – Fungal to plant biomass ratio estimation in WT and $\Delta DsNps3$ D. septosporum infected lesions (Section 3.2.3.6)

From *D. septosporum* infected lesions, qPCR analysis was done as explained in Section 2.12.1.2 to compare fungal to plant biomass ratios. The qPCR results for the summary shown in Section 3.2.3.6 are presented here. Comparisons were done separately for each technical replicate (eg. WTa-KOa, WTb-KOb, and WTC-KOc). Statistically significant difference analyses between WT and $\Delta DsNps3$ fungal to plant biomass ratios were done using t-tests by comparing WT-a-c and KOa-c for each of the samples P1, P2 and P4 separately.

Table A11. qPCR results for biomass estimation data

<table>
<thead>
<tr>
<th>Sample Name</th>
<th>Targets</th>
<th>References</th>
<th>Ct-Ref $^b$</th>
<th>Ct-Target $^c$</th>
<th>Target/Ref $^d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1-WT-a</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>24.23152289</td>
<td>29.28161173</td>
<td>0.00703</td>
</tr>
<tr>
<td>P1-WT-b</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>24.21817276</td>
<td>29.92672079</td>
<td>0.0115</td>
</tr>
<tr>
<td>P1-WT-c</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>24.69694141</td>
<td>29.13639162</td>
<td>0.00454</td>
</tr>
<tr>
<td>P1-KO2-a</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>26.59887473</td>
<td>29.70376677</td>
<td>0.0018</td>
</tr>
<tr>
<td>P1-KO2-b</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>25.99563569</td>
<td>29.46163371</td>
<td>0.0023</td>
</tr>
<tr>
<td>P1-KO2-c</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>26.55096167</td>
<td>29.72437025</td>
<td>0.00189</td>
</tr>
<tr>
<td>P2-WT-a</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>29.79491357</td>
<td>31.04160674</td>
<td>0.000503</td>
</tr>
<tr>
<td>P2-WT-b</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>29.79129381</td>
<td>30.82535338</td>
<td>0.000429</td>
</tr>
<tr>
<td>P2-WT-c</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>29.49230872</td>
<td>31.17737196</td>
<td>0.000689</td>
</tr>
<tr>
<td>P2-KO2-a</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>29.73023763</td>
<td>27.61477978</td>
<td>0.0000413</td>
</tr>
<tr>
<td>P2-KO2-b</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>29.23093224</td>
<td>27.46277611</td>
<td>0.0000526</td>
</tr>
<tr>
<td>P2-KO2-c</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>29.33066786</td>
<td>27.29554913</td>
<td>0.0000433</td>
</tr>
<tr>
<td>P4-WT-a</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>20.45181109</td>
<td>28.80877503</td>
<td>0.0722</td>
</tr>
<tr>
<td>P4-WT-b</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>20.53280436</td>
<td>28.90590667</td>
<td>0.0733</td>
</tr>
<tr>
<td>P4-WT-c</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>20.67696928</td>
<td>28.72024329</td>
<td>0.0577</td>
</tr>
<tr>
<td>P4-KO2-a</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>24.17974228</td>
<td>28.7162744</td>
<td>0.0048</td>
</tr>
<tr>
<td>P4-KO2-b</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>24.32047766</td>
<td>28.93406314</td>
<td>0.0051</td>
</tr>
<tr>
<td>P4-KO2-c</td>
<td><em>Afj</em></td>
<td><em>CAD</em></td>
<td>24.07775218</td>
<td>28.84135497</td>
<td>0.00566</td>
</tr>
</tbody>
</table>

$^a$ Sample name indicates P1-P4 plant genotype; WT/KO2 Inoculated wild type or $\Delta DsNps3$ D. septosporum strains; a-c are technical replicates.

$^b$ Cycle threshold (Ct) value of *CAD* amplification

$^c$ Cycle threshold (Ct) value of *Afj* amplification

$^d$ Fungal to plant biomass ratio based on Ct-Ref and Ct-Target.
Appendix 13 – Gene model confirmation of *DsPks1* (Section 3.3.1.2)

An amino acid alignment of DsPks1 with its best BlastP hits is presented in the CD (Appendix 21d). RNA-seq reads that were mapped onto the JGI gene model were used in order to confirm the *DsPks1* gene model (Section 3.2.3.2). The gene model was confirmed by continuous RNA-seq reads from start to end of *DsPks1*.

**Figure A13.** *DsPks1* RNA-seq reads that were mapped onto the JGI gene model. Pink rows at top are *DsPks1* RNA-seq reads at late, middle and early stages (*in planta*) and in culture (from Bradshaw et al, 2016). The numbers (eg 57.00) are RPMK values. The bold red row is predicted coding sequence (from start codon to stop codon) based on the gene model from JGI. Thin red rows indicate untranslated regions (UTRs). Green bars represent the ESTs originally used for gene annotation. PMMG and Rich media (DSM, V8, DM, PD) indicate the types of media used for growing fungal mycelia before RNA extraction. Clusters: indicate 454 EST sequences of RNA assembled with Newbler to create isotigs that were then aligned against the genomic assembly using the BLAT tool. ESTs: EST/cDNA sequences aligned against the genomic assembly using the BLAT tool. The bottom two green bars indicate the combination of the two “clusters” or “ESTs”.
Appendix 14 – Initial phylogenetic analyses of DsPks2 (Section 3.3.2.1)

Initial phylogenetic analyses were done on DsPks2 and its best BlastP hits using full protein, KS, and KS+AT sequences. As a result of all three analyses, sequences of *D. septosporum* clustered with *Aspergillus* spp., suggesting a possible horizontal gene transfer (HGT). The initial phylogenetic analyses (shown below in Figures A13a-c) were done in 2012, and therefore were based on the genome sequences available in 2012. However, detailed phylogenetic analyses done later showed that there was no HGT between *DsPks2* and its homologs (see Section 3.3.2.1).

![Phylogenetic tree of DsPKS2 based on full protein alignments.](image)

Figure A14a. Phylogenetic tree of DsPKS2 based on full protein alignments. The species indicated with red font are Dothideomycetes which have genome sequences on the Joint Genome Institute (JGI) website. The ones with green font are the Dothideomycetes which were not present on JGI, but were among the best hits as a result of BlastP searches of DsPks2 against Dothideomycetes in the NCBI database. An additional BlastP of DsPks2 against *Aspergillus spp.* was performed and the best hits were added to the alignment (black font).
Figure A14b. Phylogenetic tree of DsPKS2 based on KS domain alignments. The species indicated with red font are Dothideomycetes which have genome sequences on the Joint Genome Institute (JGI) website. The ones with green font are the Dothideomycetes which were not present on JGI, but were among the best hits as a result of BlastP searches of DsPks2 against Dothideomycetes in the NCBI database. An additional BlastP of DsPks2 against *Aspergillus spp.* was performed and the best hits were added to the alignment (black font).
Figure A14c. Phylogenetic tree of DsPKS2 based on KS+AT domain alignments. The species indicated with red font are Dothideomycetes which have genome sequences on the Joint Genome Institute (JGI) website. The ones with green font are the Dothideomycetes which were not present on JGI, but were among the best hits as a result of BlastP searches of DsPks2 against Dothideomycetes in the NCBI database. An additional BlastP of DsPks2 against *Aspergillus* spp. was performed and the best hits were added to the alignment (black font).
Appendix 15 – Gene model confirmation of DsPks2 (Section 3.3.2.2)

An amino acid alignment of DsPks2 with its best BlastP hits is presented in the CD (Appendix 21e). RNA-seq reads that were mapped onto the JGI gene model were used to confirm the DsPks2 gene model (Section 3.3.2.2). All 10 introns in the DsPks2 gene model were confirmed by RNA-seq data.

Figure A15. DsPks2 RNA-seq reads that were mapped onto the JGI gene model. Pink rows at top are DsPks1 RNA-seq reads at late, middle and early stages (in planta) and in culture (from Bradshaw et al, 2016). The numbers (e.g. 55.00) are RPMK values. The bold red row is predicted coding sequence (from start codon to stop codon) based on the gene model from JGI. Thin red rows indicate untranslated regions (UTRs). Green bars represent the ESTs originally used for gene annotation. PMMG and Rich media (DSM, V8, DM, PD) indicate the types of media used for growing fungal mycelia before RNA extraction. Clusters: indicate 454 EST sequences of RNA assembled with Newbler to create isotigs that were then aligned against the genomic assembly using the BLAT tool. ESTs: EST/cDNA sequences aligned against the genomic assembly using the BLAT tool. The bottom two green bars indicate the combination of the two “clusters” or “ESTs”.
Appendix 16 – *DsPks2*-GATEWAY gene knockout construct (Section 3.3.2.5)

The gene knock out of *DsPks2* was done using the plasmid construct shown below. This construct was prepared as in Section 2.7.1 and used in attempts to generate Δ*DsPks2* mutants using protoplast transformation as in Section 2.9.

**Figure A16.** *DsPks2*-GATEWAY KO construct. Hph-GFP shows the GFP-tagged hygromycin resistance marker. *att* sites, required for recombination in the Gateway system, at the borders of flanking regions are shown. Sizes of the whole plasmid, flanking regions, and hph-GFP are shown in brackets in bp. Primer pairs used in PCR amplification of 5' and 3' flanking sites were PKS2-73814-KO5F/PKS2-73814-KO5R and PKS2-73814-KO3F/PKS2-73814-KO3R, respectively. Primers PKS2-73814-KO5F/PKS2-73814-KO3R were used to confirm plasmid generation. Plasmid was generated based on Multisite Gateway® Three-Fragment Vector Construction Kit manual, Version G, 8 September 2008.
Appendix 17 – DsPks2 gene knockout attempts (Section 3.3.2.5)

A gene knockout of DsPks2 was attempted using the KO construct prepared as in Section 2.7.1 and transformation of D. septosporum as in Section 2.9. However, after three steps of purification the candidates still contained wild type and ΔDsPks2 gDNAs as shown by the PCR and Southern analyses below.

b)
Figure A17. PCR and Southern hybridization of $\Delta DsPks2$ candidates.

- **a)** Southern hybridization for wild type and the single $\Delta DsPks2$ candidates. SacI and EcoRV digested wild type and candidate $\Delta DsPks2$ gDNAs were hybridised to the probe shown in (c). Expected fragment sizes were 3251 bp for wt and 9303 bp for $\Delta DsPks2$ in EcoRV digested gDNAs. Expected fragment sizes for SacI digested gDNAs were 2473 bp for wt and two fragments of 629 bp and 7951 bp for $\Delta DsPks2$. The results showed that the KO candidate has both wild type and $\Delta DsPks2$, and an extra band in SacI digested gDNA could be because of a double integration.

- **b)** In attempts to obtain transformants lacking the wild type copy of the gene, six colonies were sub-cultured from the $\Delta DsPks2$ candidate by single spore purification. DNA from these was then used for PCR analysis using primers Pks2-KS-FP and D. septo Pks2Reverse (806 bp) that amplified wild type but not $\Delta DsPks2$ to confirm successful purification (wt check). Integration of the knockout construct at the $DsPks2$ locus was confirmed at both 5’ and 3’ ends of the using primer pairs Pks2-Fp-screen / hph Fwd2 (1607 bp) and GFP-Fwd / Pks2-Rp-screen (1647 bp) as shown in (c). wt: Wild type. NTC: No template control. The results showed that the KO candidates still contained both wild type and $\Delta DsPks2$ gDNAs. Inconsistency of the band strengths was probably due to the miniprep DNA extraction method that was used for rapid collection of DNA that might have extracted varying concentrations of templates or left trace amounts of chemicals that might affect PCR efficiency. Fungal internal transcribed spacer sequences (ITS) were amplified for all templates to confirm they were of sufficient quality and quantity for PCR amplification.

- **c)** Schematic representation of $DsPks2$ in wild type and $\Delta DsPks2$. The positions of SacI and EcoRV restriction enzyme recognition sites, primer binding positions, flanking regions, knock out target region in wild type (KS and AT domains) and gfp-tagged hygromycin resistance gene ($hph-gfp$) are shown.
Appendix 18 – Gene model confirmation of *DsPks3* (Section 3.3.3.1)

An amino acid alignment of DsPks3 with its best BlastP hits is presented in the CD (Appendix 21f). RNA-seq reads that were mapped onto the JGI gene model were used in an attempt to confirm the *DsPks3* gene model (Section 3.3.3.1). According to the in culture expression of *DsPks3*, only one of the introns was correctly annotated, intron 7 reading from the right.

![Figure A18a. *DsPks3* RNA-seq reads that were mapped onto the JGI gene model.](image)

Pink rows at top are *DsPks3* RNA-seq reads at late, middle and early stages (*in planta*) and in culture (from Bradshaw et al, 2016). The numbers (eg 5.50) are RPKM values. The bold red row is predicted coding sequence (from start codon to stop codon) based on the gene model from JGI. Thin red rows indicate untranslated regions (UTRs). Green bars represent the ESTs originally used for gene annotation. PMMG and Rich media (DSM, V8, DM, PD) indicate the types of media used for growing fungal mycelia before RNA extraction. Clusters: indicate 454 EST sequences of RNA assembled with Newbler to create isotigs that were then aligned against the genomic assembly using the BLAT tool. ESTs: EST/cDNA sequences aligned against the genomic assembly using the BLAT tool. The bottom two green bars indicate the combination of the two “clusters” or “ESTs”. Asterisk indicates the only intron (intron 7 reading from the right) that was correctly annotated.
When the first 6 introns are removed from \textit{DsPks3} nucleotide sequence and the resulting sequence is translated, many stop codons were found in all reading frames, suggesting that \textit{DsPks3} is pseudogenized.

Figure A18b. Three-frame translation of \textit{DsPks3} after removing first 6 introns. Introns 7, 8, and 9 are indicated with black arrows. Internal stop codons are annotated with bold blue lines. a) Reading frame 1, b) Reading frame 2, and c) Reading frame 3. Internal stop codons after intron 7 are not shown.
Appendix 19 – Gene model confirmation of *DsHps1* (Section 3.4.1.1)

An amino acid alignment of DsHps1 with its best BlastP hits is presented in the CD (Appendix 21g). Mapped RNA-seq were used to confirm the JGI *DsHps1* gene model (Section 3.4.1.1). Continuous RNA-seq reads on *DsHps1* coding sequence supported the gene model. An overlapping gene in the position of 3’ UTR suggests that this UTR may be misannotated.

**Figure A19.** *DsHps1* RNA-seq reads that were mapped onto the JGI gene model. Pink rows at top are *DsHps1* RNA-seq reads at late, middle and early stages (*in planta*) and in culture (from Bradshaw et al, 2016). The numbers (eg 6.00) are RPMK values. The bold red row is predicted coding sequence (from start codon to stop codon) based on the gene model from JGI. Thin red rows indicate untranslated regions (UTRs). A neighboring open reading frame overlapping with 3’ UTR of DsHps1 was also shown with a short, bold, and red row. Green bars represent the ESTs originally used for gene annotation. PMMG and Rich media (DSM, V8, DM, PD) indicate the types of media used for growing fungal mycelia before RNA extraction. Clusters: indicate 454 EST sequences of RNA assembled with Newbler to create isotigs that were then aligned against the genomic assembly using the BLAT tool. ESTs: EST/cDNA sequences aligned against the genomic assembly using the BLAT tool. The bottom two green bars indicate the combination of the two “clusters” or “ESTs”.
Appendix 20 – Gene model confirmation of DsHps2 (Section 3.4.2.1)

An amino acid alignment of DsHps1 with its best BlastP hits is presented in the CD (Appendix 21h). RNA-seq reads mapped onto the JGI gene model were used to confirm the DsHps2 gene model (Section 3.4.2.1). Combination of the reads from transcriptome analysis and ESTs used for gene annotation confirms all of the introns in the DsHps2 gene model.

Figure A20. DsHps2 RNA-seq reads that were mapped onto the JGI gene model. Pink rows at top are DsHps1 RNA-seq reads at late, middle and early stages (in planta) and in culture (from Bradshaw et al, 2016). The numbers (e.g. 9.00) are RPMK values. The bold red row is predicted coding sequence (from start codon to stop codon) based on the gene model from JGI. Thin red rows indicate untranslated regions (UTRs). Green bars represent the ESTs originally used for gene annotation. PMMG and Rich media (DSM, V8, DM, PD) indicate the types of media used for growing fungal mycelia before RNA extraction. Clusters: indicate 454 EST sequences of RNA assembled with Newbler to create isotigs that were then aligned against the genomic assembly using the BLAT tool. ESTs: EST/cDNA sequences aligned against the genomic assembly using the BLAT tool. The bottom two green bars indicate the combination of the two “clusters” or “ESTs”.
Appendix 23 – Whole genome sequences used in the phylogenetic analyses

All genome sequences used in the phylogenetic analyses are obtained from JGI (http://genome.jgi.doe.gov/programs/fungi/index.jsf). Abbreviations of the species and strain names used in the phylogenetic trees were same as in JGI. The list below shows the genomes used as “abbreviation: species and strain name, and (reference) or principal investigator of the genome sequencing project.

Aaor1: *Aeosphaeria arxii* CBS 175.79 v1.0 – Pedro Crous: p.crous@lbs.knaw.nl
Acacid1: *Acidothrix acidophilia* CBS 136259 v1.0 (Hujslová et al., 2014)
Acri1_iso: *Acidomyces richmondensis* BFW (Mosier et al., 2016)
Acri1_meta: *Acidomyces richmondensis* BFW (Mosier et al., 2016)
Acral1_c3a: *Acremonium alcalophilum* v1.0 – Adrian Tsang: tsang@gene.concordia.ca
Acral2: *Acremonium alcalophilum* v2.0 – Adrian Tsang: tsang@gene.concordia.ca
Arch1: *Acremonium strictum* DS1bioAY4a v1.0 – Colleen Hansel: chansel@whoi.edu
Altai: *Alternaria alternata* SRC1K2v1.0 (Zeiner et al., 2016)
Altbr1: *Alternaria brassicicola* – Christopher Lawrence: lawrence@vbi.vt.edu
Amnli1: *Amniculicola lignicola* CBS 123094 v1.0 – Pedro Crous: p.crous@lbs.knaw.nl
Ampqui1: *Ampelomyces quisqualis* HMLAC05119 v1.0 – Chen Liang: syliangchen@163.com
Antav1: *Anthostoma avocetta* NRR1 3190 v1.0 – Kerry O'Donnell: kerry.odonnell@ars.usda.gov
Apimo1: *Apiospora montagnei* NRRL 25634 v1.0 – Pedro Crous: p.crous@lbs.knaw.nl
Aplpr1: *Aplosporella prunicola* CBS 121.167 v1.0 – Pedro Crous: p.crous@lbs.knaw.nl
Artbe1: *Arthroderma benhamiae* CBS 112371 (Burmester et al., 2011)
Aspac1: *Aspergillus aculeatus* ATCC16872 v1.1 – Scott Baker: scott.baker@pnl.gov
Aspbr1: *Aspergillus brasiliensis* v1.0 (de Vries et al., 2017)
Aspca: *Aspergillus carbonarius* v1.0 – Scott Baker: scott.baker@pnl.gov
Aspcam1: *Aspergillus campestris* IBT 28561 v1.0 – Scott Baker: scott.baker@pnl.gov
Aspfl1: *Aspergillus flavus* NRRL3357 (Arnaud et al., 2012)
Aspfo1: *Aspergillus luchuensis* CBS 106.47 v1.0 (de Vries et al., 2017)
Aspgl1: *Aspergillus glaucus* v1.0 (de Vries et al., 2017)
Aspni: *Aspergillus niger* ATCC 1015 v3.0 (Andersen et al., 2011)
Aspni: *Aspergillus niger* ATCC 1015 v4.0 (Andersen et al., 2011)
Aspni: *Aspergillus niger* NRRL3 v1.0 (de Vries et al., 2017)
Aspso1: *Aspergillus ochraceoroseus* IBT 24754 v1.0 – Scott Baker: scott.baker@pnl.gov
Aspste1: *Aspergillus steynii* IBT 23096 v1.0 – Scott Baker: scott.baker@pnl.gov
Aspsy1: *Aspergillus sydowii* CBS 593.65 v1.0 (de Vries et al., 2017)
Aspte1: *Aspergillus terreus* NIH 2624 (Arnaud et al., 2012)
Asptu1: *Aspergillus tubingensis* CBS 593.65 v1.0 (de Vries et al., 2017)
Aulhe2: *Aulographum hederae* v2.0 – Pedro Crous: p.crous@lbs.knaw.nl
Aurpu_var_mel1: *Aureobasidium pullulans* var. *melanogenum* CBS 110374 (Correct taxonomic name: *Aureobasidium melanogenum*) (Gostinčar et al., 2014)
Aurpu_var_nam1: *Aureobasidium pullulans* var. *namibiae* CBS 147.97 (Correct taxonomic name: *Aureobasidium namibiae*) (Gostinčar et al., 2014)
Aurpu_var_pull1: *Aureobasidium pullulans* var. *pullulans* EXF-150 (Correct taxonomic name: *Aureobasidium pullulans*) (Gostinčar et al., 2014)
Bauco1: *Baudoinia compniacensis* UAMH 10762 (4089826) v1.0 (Ohm et al., 2012)
Gymau1: Gymnascella aurantiaca v1.0 – Kerry O’Donnell: kerry.odonnell@ars.usda.gov
Gymci1_1: Gymnascella citrina v1.1 – Kerry O’Donnell: kerry.odonnell@ars.usda.gov
Horac1: Hortaea acidophila CBS 113389 v1.0 – Jon Karl Magnuson: jon.magnuson@pnnl.gov
HypC4A_1: Hypoxylon sp. C1-4A v1.0 (Wu et al., 2017)
HypD0275.1: Hypoxylon sp. C027-5 v1.0 (Wu et al., 2017)
HypEC38_1: Hypoxylon sp. EC38 v1.0 – Blake Simmons: basimmo@sandia.gov
Hyspu1: Hysterium pulicare (Ohm et al., 2012)
Ilysp1: Ilyonectria sp. v1.0 – Francis Martin: fmartin@nancy.inra.fr
Karrh1: Karstenula rhodostoma CBS 690.94 v1.0 – Pedro Crous: p.crous@cbs.knaw.nl
Khuory1: Khuskia oryzae ATCC 28132 v1.0 – Jon Karl Magnuson: jon.magnuson@pnnl.gov
Lenfl1: Lentithecium fluviatile v1.0 – Pedro Crous: p.crous@cbs.knaw.nl
Lepmu1: Leptosphaeria maculans (Rouxel et al., 2011)
Leppa1: Lepidopterella palustris v1.0 (Peter et al., 2016)
Linin1: Lindigomyces ingoldianus ATCC 200398 v1.0 – Jon Karl Magnuson: Jon.Magnuson@pnnl.gov
Linrh1: Lineolata rhizophorae ATCC 116597 v1.0 – Patrik Inderbitzin: prin@ucdavis.edu
Linth1: Lindra thalassiae JK4322 v1.0 – Derek Johnson: johnsde4@science.oregonstate.edu
Lizem1: Lizonia empirigonia CBS 542.76 v1.0 – Pedro Crous: p.crous@cbs.knaw.nl
Lolmi1: Lollipopaia minuta P26; CBS 116597 v1.0 – Patrik Inderbitzin: prin@ucdavis.edu
Lopma1: Lophiostoma macrostomum v1.0 – Pedro Crous: p.crous@cbs.knaw.nl
Lopmy1: Lophium mytilinum CBS 269.34 v1.0 – Pedro Crous: p.crous@cbs.knaw.nl
Macan1: Macroventuria anomochaeta CBS 525.71 v1.0 – Pedro Crous: p.crous@cbs.knaw.nl
Maggr1: Magnaporthopsis grisea v1.0 (Dean et al., 2005)
Maseb1: Massarina eburnea CBS 473.64 v1.0 – Pedro Crous: p.crous@cbs.knaw.nl
Melpu1: Melanomma pulvis-pyrius v1.0 – Pedro Crous: p.crous@cbs.knaw.nl
Melsp1: Melanconium sp. NRRL 54901 v1.0 – Joseph Spatafora: spatafoj@science.oregonstate.edu
Melti1: Microdochium bolleyi J235TASD1 v1.0 (David et al., 2016)
Myche1: Myceliophthora heterothallica CBS 203.75 v1.0 – Don Natvig: dnatvig@gmail.com
Myrdu1: Myriangium duriaei CBS 260.36 v1.0 – Pedro Crous: p.crous@cbs.knaw.nl
Mytre1: Mytilinidion resinicola CBS 304.34 v1.0 – Jon Karl Magnuson: Jon.Magnuson@pnnl.gov
Nieex1: Niesslia exilis CBS 358.70 v1.0 – Kathryn Bushley: kbushley@umn.edu
Myctr1: Microascus trigonosporus CBS 218.31 v1.0 – Kathryn Bushley: kbushley@umn.edu
Ne cata2: Nectria haematococca v2.0 (Coleman et al., 2009)
Neucr1: Neurospora crassa OR74A v1.0 (Galanag et al., 2003)
Neucr2: Neurospora crassa OR74A v2.0 (Galanag et al., 2003)
Neucr_trp3_1: Neurospora crassa FGSC 73 trp-3 v1.0 (Baker et al., 2015)
Neud1: Neurospora discreta FGSC B579 mat A – John Taylor: jtaylor@nature.berkeley.edu
Neute1: Neurospora tetrasperma FGSC 2508 mat A – John Taylor: jtaylor@nature.berkeley.edu
Neute_matA2: Neurospora tetrasperma FGSC 2508 mat A v2.0 (Ellison et al., 2011)
Neute_mat_a1: Neurospora tetrasperma FGSC 2509 mat a v1.0 (Ellison et al., 2011)
Nieex1: Niesslia exilis CBS 358.70 v1.0 – Kathryn Bushley: kbushley@umn.edu
Opohl1: Ophiobolus disseminans CBS 113818 v1.0 – Pedro Crous: p.crous@cbs.knaw.nl
Opophc1: Ophiostoma piceae UAMH 11346 (Haridas et al., 2013)
Parspl1: Paraconiothyrium sporulosum AP3s5-jAC2a v1.0 (Zeiner et al., 2016)
Patat1: Patellaria atrata v1.0 – Pedro Crous: p.crous@cbs.knaw.nl
Penac1: Talaromyces aculeatus ATCC 10409 v1.0 – Keith Seifert: keith.seifert@agr.gc.ca
Penbi1: Penicillium brevicompactum ATCC 20851 v1.0 – Dave Greenshields: dvgs@novozymes.com
Penbr2: Penicillium brevicompactum 1011305 v2.0 – Dave Greenshields: dvgs@novozymes.com
PenbrAgRF18_1: Penicillium brevicompactum AgRF18 v1.0 – Dave Greenshields: dvgs@novozymes.com
Penca1: Penicillium canescens ATCC 10419 v1.0 – Dave Greenshields: dvgs@novozymes.com
Pench1: Penicillium chrysogenum v1.0 – Igor Grigoriev: IGVGrigoriev@lbl.gov
Penex1: Penicillium expansum ATCC 24692 v1.0 – Dave Greenshields: dvgs@novozymes.com
Penfe1: Penicillium fellutanum ATCC 48694 v1.0 – Dave Greenshields: dvgs@novozymes.com
Pengl1: Penicillium glabrum DAOM 239074 v1.0 – Dave Greenshields: dvgs@novozymes.com
Penja1: Penicillium janthinellum ATCC 10455 v1.0 – Dave Greenshields: dvgs@novozymes.com
Penla1: Penicillium lanascoeuruleum ATCC 48919 v1.0 – Dave Greenshields: dvgs@novozymes.com
Penna1: Penicillium raiistricki ATCC 10490 v1.0 – Dave Greenshields: dvgs@novozymes.com
Perm1: Periconia macrospora DSE2036 v1.0 – Gabor M. Kovacs: gmkovacs@elte.hu
PhaPMI808: Phaeosphaeraceae sp. PMI_808 v1.0 – Gregory Bonito: binoito@msu.edu
Photr1: Phoma tracheiphila ipt5 v1.0 – David Ezra: Dezra@volcaniagri.gov.il
Phycit1: Phyllosticta citriasiana v1.0 – Pedro Crous: p.crous@ubs.knaw.nl
Pieho1: Plectrospora cucumerina DS2psM2a2 v1.0 – Colleen Hansel: chansel@who.edu
Plesi1: Pleosporina siparia v1.0 – Pedro Crous: p.crous@ubs.knaw.nl
Podcru1: Podospora curvicolla TEP21a v1.0 – Kabir Peay: kpeay@stanford.edu
Polci1: Polychaeton citri v1.0 – Pedro Crous: p.crous@ubs.knaw.nl
Poliuf1: Polyplephosphaeria fusca CBS 125425 v1.0 – Pedro Crous: p.crous@ubs.knaw.nl
Psehyl1: Pseudovirgaria hyperparasitica CBS 121739 v1.0 – Pedro Crous: p.crous@ubs.knaw.nl
Pseve2: Pseudomassariella vexata CBS 129021 v1.0 – Jon Karl Magnuson: Jon.Magnuson@pnnl.gov
Pyrs1: Pyrenoacheta sp. DS3sAY3a v1.0 (Zeiner et al., 2016)
Pytr1: Pyrenophora tritici-repentis (Manning et al., 2013)
Pytrt1: Pyrenophora teres f. teres (Blwwood et al., 2010)
Rhili: Rhizodiscina lignyota CBS 133067 v1.0 – Pedro Crous: p.crous@ubs.knaw.nl
Rhryu1: Rhytidhysteran rufulum (Ohm et al, 2012)
Sacpr1: Saccharata proteae CBS 121410 v1.0 – Pedro Crous: p.crous@ubs.knaw.nl
Sepmu1: Septoria musiva SO2202 v1.0 (Ohm et al. 2012; Dhillon et al. 2015)
Setho1: Setomelanomma holmii CBS 110217 v1.0 – Pedro Crous: p.crous@ubs.knaw.nl
Settu1: Setosphaeria turcica Et28A v1.0 (Ohm et al., 2012; Condon et al., 2013)
Settur1: Setosphaeria turcica NY001 v1.0 – B. Gillian Turgeon: bgt1@cornell.edu
Sodali: Sodiomyces alkalinus v1.0 – Alexey Grum Grzhimaylo: alexey.grumgrzhimaylo@wur.nl
Spofi1: Sporormia fimetaria v1.0 – Pedro Crous: p.crous@ubs.knaw.nl
Spoth1: Sporotrichum thermophile v1.0 – Adrian Tsang: tsang@gene.concordia.ca
Spoth2: Myceliophthora thermophila (Sporotrichum thermophile) v2.0 (Berka et al., 2011)
Stano1: Stagonospora nodorum SN15 (Correct taxonomic name: Parastagonospora nodorum) (Hane et al., 2007)
Staspr1: Stagonospora sp. SRC1lsM3a v1.0 (Zeiner et al., 2016)
Ternu1: Ternatodactylus nubilosa CBS 116005 v1.0 – Pedro Crous: p.crous@ubs.knaw.nl
Theau1: Thermoascus aurantiacus v1.0 – Steven Singer: SWSinger@lbl.gov
Thian1: Thielavia antarctica CBS 123565 v1.0 – Adrian Tsang: tsang@gene.concordia.ca
Thiap1: Thielavia appendiculata CBS 731.68 v1.0 – Adrian Tsang: tsang@gene.concordia.ca
Thiar1: Thielavia arenaria CBS 508.74 v1.0 – Adrian Tsang: tsang@gene.concordia.ca
Thily1: Thielavia hyrcaniae CBS 757.83 v1.0 – Adrian Tsang: tsang@gene.concordia.ca
Thit1: Thielavia terrestris v1.0 (Berka et al., 2011)
Thite2: Thielavia terrestris v2.0 (Berka et al., 2011)
ThoPMI491_1: Thozetella sp. PMI_491 v2.0 – Gregory Bonito: binoito@msu.edu
Torra1: Torpedospora radiata jKs526c v1.0 – Derek Johnson: johnsde4@science.oregonstate.edu
Toffu1: Totiha fuscella CBS 130226 v1.0 – Pedro Crous: p.crous@ubs.knaw.nl
Trepe1: Trematosphaeria pertusa CBS 122368 v1.0 – Pedro Crous: p.crous@ubs.knaw.nl
Trias1: Trichoderma asperellum CBS 433.97 v1.0 – Igor Grigoriev: IGVGrigoriev@lbl.gov
Triasp1: Trichoderma asperellum TR356 v1.0 – Igor Grigoriev: IGVGrigoriev@lbl.gov
Triat2: Trichoderma atroviride v2.0 – Scott Baker: scott.baker@pnl.gov
Trib1: Trichoderma atroviride v1.0 – Christian Kubicek: cubikeck@mail.zserv.tuwien.ac.at
Trici1: Trichoderma citrinoviride v1.0 – Christian Kubicek: cubikeck@mail.zserv.tuwien.ac.at
Trici4: Trichoderma citrinoviride TUCIM 6016 v4.0 – Igor Grigoriev: IGVGrigoriev@lbl.gov
Triha1: Trichoderma harzianum CBS 226.95 v1.0 – Igor Grigoriev: IGVGrigoriev@lbl.gov
Trilo1: Trichoderma longibrachiatum ATCC 18648 v1.0 – Christian Kubicek: cubikeck@mail.zserv.tuwien.ac.at
Trilo3: Trichoderma longibrachiatum ATCC 18648 v3.0 – Igor Grigoriev: IGVGrigoriev@lbl.gov
Trire2: *Trichoderma reesei* v2.0 (Martinez et al., 2008)
TrireRUTC30_1: *Trichoderma reesei* RUT C-30 v1.0 (Koike et al., 2013)
TriviGv29_8_2: *Trichoderma virens* Gv29-8 v2.0 – Charles M. Kenerley: c-kenerley@tamu.edu
Tryvi1: *Trypethelium eluteriae* v1.0 – Stephen Goodwin: sgoodwin@purdue.edu
Valla1: *Valetoniellopsis laxa* CBS 191.97 v1.0 – Kathryn Bushley: kbushley@umn.edu
Verda1: *Verticillium dahliae* v1.0 (Klosterman et al., 2011)
Veren1: *Verruculina enalia* CBS 304.66 v1.0 – Pedro Crous: p.crous@cbs.knaw.nl
Wesor1: *Westerdykella ornata* CBS 379.55 v1.0 – Pedro Crous: p.crous@cbs.knaw.nl
Zasce1: *Zasmidium cellare* ATCC 36951 v1.0 – Stephen Goodwin: sgoodwin@purdue.edu
Zoprh1: *Zopfia rhizophila* v1.0 – Pedro Crous: p.crous@cbs.knaw.nl
References of the whole genome sequences


de Wit, P.J.G.M., A. van der Burgt, B. Okmen, I. Stergiopoulos, K.A. Abd-Elsalam et al. (2012). The genomes of the fungal plant pathogens *Cladosporium fulvum* and
*Dothistroma septosporum* reveal adaptation to different hosts and lifestyles but also signatures of common ancestry. *PloS Genetics*, 8, e1003088.


Haridas, S., Wang, Y., Lim, L., Massoumi Alamouti, S., Jackman, S., Docking, R. et al. (2013). The genome and transcriptome of the pine saprophyte *Ophiostoma piceae*, and a comparison with the bark beetle-associated pine pathogen *Grosmannia clavigera*. *BMC Genomics*, 14, 373.


References


