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

List of Tables

List of Figures

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
ddH_2O	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

dpi	Days post-inoculation
dS	Synonymous mutation number
DSM	Dothistroma sporulation medium
ER	Enoylreductase
EtAc	Ethyl acetate
EtBr	Ethidium bromide
EtOH	Ethanol
ETS	Effector-triggered susceptibility
F (domain)	Formylation
FAD	Flavin adenine dinucleotide
FAS	Fatty acid synthase
FDR	False discovery rate
GFP	Green fluorescent protein
HPLC	High performance liquid chromatography
Hph	Hygromycin B phosphotransferase
HPS	Hybrid polyketide - nonribosomal peptide synthetase
HR	Hypersensitive response
HR-PKS	Highly-reducing PKS
HST	Host-specific toxin
kb	Kilo base pairs
KO	Knockout mutant
KR	Ketoreductase
KS	Keto-synthase
LAS	Leica application suite

LB	Luria broth
MAFFT	Multiple alignment fast Fourier transform
MCO	Multicopper oxidase
MeOH	Methanol
MFS (transporter)	Major facilitator superfamily
ML	Maximum likelihood
MMIC	The Manawatu Microscopy & Imaging Centre
MS	Mass spectrometry
NaPDoS	Natural product domain seeker
NCBI	National center for biotechnology information
NHST	Non-host-specific toxin
NM	N-methyltransferase
NMR	Nuclear magnetic resonance
Non-norm	Non-normalized
Norm.	Normalized
NR-PKS	Non-reducing PKS
NRP	Nonribosomal peptide
NRPS/NPS	Nonribosomal peptide synthetase
Ox (domain)	Oxidation
PAMP/MAMP	Pathogen/microbe-associated molecular pattern
PKS	Polyketide synthase
PMMG	Pine needle minimal medium + glucose
PR-PKS	Partially reducing PKS
PT	Product template

PTI	PAMP-triggered immunity
R (domain)	Reduction
R (gene/protein)	Resistance
Rf	Retention factor
RPMK	Fungal reads per million per kilobase
SAM	S-adenosyl-L-methionine
SAT	Starter unit acyl-carrier protein transacylase
SBSPKS	Structure based sequence analysis of polyketide synthases
SCD	Scytalone dehydratase
SLR	Sitewise likelihood-ratio
SM	Secondary metabolite
ST	Sterigmatocystin
T	Thiolation
T3HN	1,3,8- trihydroxynaphthalene
T4HN	1,3,6,8-tetrahydroxynaphthalene
TE	Thioesterase
TLC	Thin layer chromatography
wpi	Weeks post-inoculation
WT	Wild type
µg	Microgram
µL	Microliter
µM	Micromolar