Molecular and bioinformatic analysis of the *perA* locus in *Epichloë*

This thesis is presented as a partial fulfilment of the requirements for the degree of

Master of Science (MSc)

in

Genetics

at Massey University, Palmerston North

New Zealand

Daniel Berry

2011

Abstract

Fungal endophytes of the *Epichloë* genus form largely mutualistic symbioses with cool-season grasses, systemically colonising the intercellular spaces of the host in a strictly regulated fashion. The endophyte receives protection and sustenance from the host, and in return provides benefits such as increased growth, drought resistance and protection against herbivores. Protection against herbivory is mediated through the production of bio-protective fungal secondary metabolites (SM). Examples of these SMs include lolitrem B, the causative agent of 'ryegrass staggers' in stock, and the insect feeding deterrent peramine.

The genes responsible for the production of each of these SMs are usually found clustered together in the genome, and are often closely associated with a range of transposon relics. SM gene expression occurs only when the endophyte is growing *in planta*, indicating the presence of plant-fungal signalling. This study investigated the locus structure and organisation of the gene *perA* that encodes the non-ribosomal peptide synthetase PerA, which is both essential and sufficient for production of peramine. It was found that *perA* and its flanking intergenic sequences exhibit considerable transposon-mediated variability across *Epichloë*, and that this transposon activity is likely responsible for the taxonomically discontinuous production of peramine both within and across *Epichloë spp*.

The major facilitator superfamily transporter gene *EF102* is divergently transcribed from and co-regulated with *perA* (*EF103*). Transcriptome data were used to identify transcription start sites for both genes. Comparative analysis of the intergenic sequence separating *EF102/perA* from 10 *Epichloë* isolates covering six different species refined the *perA* translation start site, and identified conserved regions in the promoters of both genes proposed to be important for regulation. A motif search identified a conserved DNA motif present multiple times in the promoters of both genes.

Deletion analysis of *EF102* revealed the gene probably does not encode a peramine transporter, as was hypothesised; however the four independent Δ*EF102* mutants exhibited a reduction in peramine production relative to wild type, resulting in an alternative hypothesis that *EF102* encodes a transporter for a PerA substrate precursor molecule such as glutamate.

Acknowledgements

Firstly I thank my supervisor Barry Scott for his guidance and support (and funding!), without which this would not have been possible.

Thanks to my collaborator Carolyn Young, who provided much of the material and DNA sequences required for this study.

Thanks to the Scottbase members (both past and present) Ruth, Gemma, Carla, Emma, Milena, Tetsuya, Yvonne, Sarah, Matt, Sanjay for their help and guidance around the lab over the last two (and a bit) years, and to Murray for his help with the bioinformatics aspects of this study. Thanks to Mike for getting me into this area in the first place.

Thanks to the lovely people at AgResearch who keep our plants watered, and especially Wayne and Wade for their advice and work on the guttation analyses.

Thanks to IMBS and all the staff (esp. Ann, Pat and Cynthia) who have helped me along the way.

Lastly thanks to Mum and Dad for their support – both emotional and fiscal**!**

Table of Contents

List of Tables

List of Figures

Common Abbreviations

A1/2	NRPS adenylation domain from modules 1 or 2
Amp	Ampicillin
ASW	Argentine stem weevil
ATG	ATG translational start codon
bp	base pairs
C1	NRPS condensation domain from module 1
CDS	Coding sequence
d	Days
dATP	Deoxyadenine triphosphate
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
ECM	Extracellular matrix
EDTA	Ethylene diamine tetra-acetic acid
EF	*Epichloë festucae*
g	Gravity
gDNA	Genomic DNA
h	Hours
HGT	Horizontal gene transfer
HPLC	High-performance liquid chromatography
Hyg	Hygromycin
Indel	Insertion or deletion
kb	Kilo base pairs
KO	Knock out
LB	Luria-Bertani medium
LB	Left border
LC	Liquid chromatography
M	Molar
M1	NRPS methylation domain from module 2
min	Minutes
MITE	Miniature inverted transposable element
mRNA	Messenger RNA
MFS	Major facilitator superfamily
MS	Mass spectrophotometer
MSA	Multiple sequence alignment
NCM	Nitrocellulose membrane
ND	Not detected
NRPS	Non-ribosomal peptide synthetase
NT	Not tested
P5C	1-pyrroline-5-carboxylate
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PD	Potato dextrose medium
PKS	Polyketide synthase
R2/Rdom	NRPS reductase domain from module 2

RB	Right border
RCF	Relative centrifugal force
RE	Restriction enzyme
RG	Regeneration medium
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RT-PCR	Reverse transcription PCR
SDS	Sodium dodecyl sulfate
SSC	Saline sodium citrate
SM	Secondary metabolite
SNP	Single nucleotide polymorphism
T1/2	NRPS thiolation domain from module 1 or 2
TBE	Tris Borate EDTA buffer
UV	Ultra-violet
WT	Wild type