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.

Identification of Transporter Genes from the Fungal Endophyte *Neotyphodium lolii*

This Thesis is presented in partial fulfilment of
the requirements for the degree of
Master of Science (MSc) in Plant Biology
at Massey University, Palmerston North
New Zealand

Deborah Sian Knox

2006

Abstract

Neotyphodium lolii is an endophytic fungus that lives in the pasture grass, *Lolium perenne*. They share a mutualistic symbiotic relationship. *N. lolii* lives out its life cycle within the plant and produces secondary metabolites, including alkaloids peramine, ergovaline and lolitrem which protect the grass from insect and animal herbivory. In fungi the biosynthetic genes of secondary metabolites are often located in gene clusters. These clusters frequently contain one or more genes that code for transporter proteins responsible for the removal of toxic products from the fungal cells. Plants produce defence compounds, including antifungals to protect themselves from colonising fungi. However endophytes are able to neutralise these host toxins, one mechanism for this is possibly by efflux through transporter channels.

The goal of this study was to identify ABC and MFS genes from *N. lolii*. These two families are the largest and most diverse of transporter families, which transport a variety of substrates, including peptides, toxins, ions and sugars across membranes. Using degenerate PCR primers designed from fungal multi-drug transporter sequences, four unique ABC gene fragments were amplified from *N. lolii*. A further two ABC sequences and two MFS gene fragments were identified in a database of *N. lolii* EST sequences.

RT-PCR was used to compare expression of isolated ABC and MFS genes in *N. lolii*, growing in culture and in infected plants. Up-regulation of transporter transcripts *in planta* could suggest a role in symbiosis. Some genes were seen to have a visibly different expression pattern from others, although all genes were strongly expressed in cultured mycelia. Gene expression in the plant host was most evident in tissues more heavily infected with endophyte. To discover possible roles for the isolated transporter genes in transporting endophyte secondary metabolites a strain distribution study was completed. Five of the putative ABC and MFS genes were compared against 12 *Epichloë* and *Neotyphodium* endophytes. Amplified PCR products in the genotypes screened produced a unique pattern of gene occurrence for each of the five transporters. This added to the characterisation of the transporter genes and showed that one gene,

gABC 4c, was the most diverse in its distribution, while another ABC gene gABC 4g was present across all genotypes.

One ABC gene (gABC 4c) plus flanking DNA was sequenced in full. Bioinformatic analyses suggested that gABC 4c may be a half sized ABC transporter gene of 2 kb with four exons. An orotate phosphoribosyltransferase was identified 2 kb upstream of the ABC transporter.

Further work will be needed to confirm that the start and stop codons of this ABC transporter have been accurately predicted, as well as to verify the putative intron/ exon boundaries identified by gene prediction programmes. The role of *N. lolii* ABC transporter gABC 4c has not been determined, however future research could focus on the nature of the substrate(s) transported, the sub-cellular location of the channel, and the effects of gene knockout or over-expression on the symbiosis between *N. lolii* and perennial ryegrass.

Acknowledgements

To my supervisors Christine Voisey and Rosie Bradshaw for all their help, suggestions and support. Thank you for all your encouragement when I had no idea what to do next and the time you gave to me from your very busy lives. Also thanks to Richard Johnson, for his input in this project and Gregory Bryan.

Thanks to the all the people in the AFT lab over the years, particularly Shalome, Damien, Linda and Charlotte for all their help and patience in explaining new techniques and ideas, and just passing on their accumulated wisdom. Especial thanks to Jen for her friendship and distractions, and making life in the cupboard more bearable.

Everyone else at AgR who has helped me out, taught me and previously given me a job, or encouragement. You all made the working environment so friendly and encouraging, you were the reason I decided to stay and do my Masters in the first place by giving me an encouraging view of research science in the real world.

Always thanks to my family, for their love and encouragement. I want to thank my parents for all their support and love and to my sister Shelles for always being there and listening.

And for those friends I am connected with via email, thanks so much for the discussions, daily life updates (sometimes up-to-the-minute) and light amusement to improve every day.

I would lastly like to thank my cat Orion, and the boys. Just for insisting by yowls or squeals that I get out of bed every morning (just to feed them of course).

Table of Contents

Chapter 1 Introduction 1

1.1 *Lolium perenne* and *Neotyphodium lolii*, mutualistic symbionts 2

1.1.1 Endophytic Interactions 2

1.1.2 Alkaloid Production; Benefits to the Host 3

1.2 Fungal Secondary Metabolites and Gene Clusters 5

1.2.1 Gene Clusters and the Production of Fungal Secondary Metabolites 5

1.2.2 Gene Clusters in Endophyte Genomes 6

1.2.2.1 Lolines 6

1.2.2.2 Lolitrems 6

1.2.2.3 Ergot Alkaloids 7

1.2.2.4 Peramine 7

1.3 ABC and MFS transporters 9

1.3.1 Identification of Transporters in Fungi 9

1.3.2 ABC Transporters 9

1.3.2.1 Structure of ABC Transporters 10

1.3.2.2 Classes of ABC Transporters 10

1.3.2.3 Mechanism of ABC Transporters 11

1.3.2.4 Roles of ABC Transporters 12

1.3.3 MFS Transporters 13

1.3.3.1 Structure of MFS Transporters 13

1.3.3.2 Classes of MFS Transporters 14

1.3.3.3 Mechanism of MFS Transporters 15

1.3.3.4 Roles of MFS Transporters 15

1.4 Identifying ABC and MFS transporter genes in *Neotyphodium lolii* 18

1.4.1 Identification and Analysis of Transporter Genes 18

1.4.2 Identification of Transporter Genes in *N. lolii* 18

1.4.3 Expression of Transporter Genes 19

1.4.4 Distribution of Endophyte Transporter Genes 20

1.5 Aims and Objectives 22

Chapter 2 Materials and Methods 23

2.1 Growth and Maintenance of Cultures 24

2.1.1 Growth of *E. coli* Cultures 24

2.1.2 Growth of Fungal Cultures 24

2.2 Nucleic Acid Extractions 26

2.2.1 Protoplast Method 26

2.2.2 High Molecular Weight Extraction of Endophyte Genomic DNA 27

2.2.3 Isolation of Plasmid DNA from *E. coli* 28

2.3 Molecular Techniques 28

2.3.1 Polymerase Chain Reaction (PCR) 28

2.3.1.1 Standard PCR 28

2.3.1.2 Degenerate PCR 29

2.3.2 Gel Electrophoresis 32

2.3.3 Gel Extraction 32

2.3.4 Nucleic Acid Concentration 32

2.3.5 Restriction Digests 32
2.3.6 Ligations 32
2.3.7 Transformations 33
2.3.8 Sequencing 33
2.3.9 Preparation of Sequencing Template DNA 34
2.4 RNA Procedures 36
2.4.1 RNA Extraction 36
2.4.2 cDNA Synthesis 37
2.4.3 Reverse Transcriptase PCR 37
2.5 Library Screening 38
2.5.1 BAC Library Screening 38
2.5.2 λ Library Screening 39
2.5.3 Plaque Lifts 39
2.6 Southern Blotting with DIG-Labelled Probes 40
2.6.1 Labelling of DNA Probes 40
2.6.2 Hybridisation of Southern Blots and λ Plaque Lifts 40
2.6.3 Stripping Membrane-Bound Probes and Reusing Probe Solutions 42
2.7 Bioinformatics 42
2.7.1 BLAST Searches 42
2.7.2 Sequence Alignments 42
2.7.3 Sequence Assembly 43
2.7.4 Gene Annotation 43

Chapter 3 Results 44
3.1 Isolation ABC and MFS Transporters from Genomic DNA of *N. lolii* 47
3.1.1 Multidrug Transporters and the *N. lolii* Genome 47
3.1.2 Fungal ABC and MFS Genes 47
3.1.3 Alignment of MFS Transporter Genes and Primer Design 48
3.1.3.1 MFS PCR Optimisation 50
3.1.3.2 Cloning of Putative PCR Fragments 51
3.1.4 Alignment of ABC Transporter Genes and Primer Design 53
3.1.4.1 ABC Alignment 1 53
3.1.4.2 ABC Alignments Two and Three 54
3.1.4.3 Sequencing of Putative ABC and MDR Gene Fragments 58
3.1.4.4 Nested PCR Primers Improves the Identification of ABC Transporters 59
3.1.4.5 Sequencing Putative ABC Transporter Genes from Nested PCR Fragments 61
3.1.5 Identification of ABC and MFS Transporters from a Database of *N. lolii* ESTs 67
3.1.6 Summary of ABC and MFS Transporter Gene Identification 70
3.1.6.1 MFS Transporters: Low Sequence Conservation 70
3.1.6.2 ABC Transporter Identification: Success in Nesting 70
3.2 Analysis and Expression of *N. lolii* ABC and MFS Transporters 73
3.2.1 Semi-Quantitative Reverse-Transcriptase PCR of putative transporter genes 73
3.2.2 Analysis of Transporter Sequence Fragments for RT-PCR Primer Design 75

3.2.3 Design and PCR Optimisation of ABC and MFS Primers for RT-PCR 79
3.2.3.1 Design and Optimisation of EST Primers 79
3.2.3.2 Genomic Primers Design and Optimisation 81
3.2.4 Reverse-Transcriptase Expression Analysis 82
3.2.4.1 RNA Extraction 82
3.2.4.2 cDNA synthesis 83
3.2.4.3 cDNA Expression of 8 Different Transporter Genes 83
3.3 Identifying Transporter Genes in the BAC Library 86
3.3.1 Screening the Pooled BAC Plate 86
3.3.2 Selecting Putative Transporters of Interest 89
3.4 Distribution of Transporters in Other Endophyte Strains 91
3.4.1 Primer Optimisation 91
3.4.2 Presence of transporters across 12 Endophyte strains 92
3.5 Sequencing and Annotation of Transporter Gene gABC 4e and Flanking DNA 97
3.5.1 Sequencing of a Lambda Phagemid Insert 97
3.5.2 Analysis of the Complete Sequence of an ABC Transporter Gene from *N. lolii* 97
3.5.3 Searching for Exons: Gene Annotation 99

Chapter Four Discussion and Conclusions 108
4.1 Identifying ABC and MFS Genes in the *N. lolii* Genome 109
4.2 Expression of ABC and MFS Genes 111
4.3 ABC and MFS Genes in the Genomes of Endophyte and Other Fungi 114
4.4 Gene Analysis 117
4.5 Conclusion 119

References 122

Appendices 131
Appendix I Media 131
Appendix II Common Buffers and Solutions 131
Appendix III Vector Maps 133
Appendix IV Additional Transporter Alignments 134
Appendix V Complete Sequence of Phagemid Insert with gABC 4e 142

List of Figures

Figure 1.1 Alkaloids of *N. lolii* 4
Figure 1.2 Structure of an ABC Transporter 10
Figure 1.3 Structure of a MFS Transporter 14
Figure 1.4 Mechanisms of Substrate Transport by MFS 15
Figure 3.1 MFS Alignment 1 49
Figure 3.2 Motifs of the MFS transporters 50
Figure 3.3 Optimisation of $MgCl_2$ Concentration and Annealing Temperature of Degenerate Primers MFS 1F/2R 51
Figure 3.4 PCR Confirmation of Initial Putative MFS Transformants 52
Figure 3.5 PCR Optimisations for Primers ABC 1F/2R 54
Figure 3.6 ABC Alignment 2 56
Figure 3.7 PCR Optimisation for Primers ABC 4F and ABC 5R 57
Figure 3.8 PCR Optimisation for Primers MDR 1F and MDR 2R 58
Figure 3.9 PCR Confirmations of ABC and MDR Transformations 59
Figure 3.10 Nested PCR step 1 60
Figure 3.11 Nested PCR step 2 61
Figure 3.12 BLASTX Alignment 62
Figure 3.13 Alignment tree containing the 88 GA 300 sequences grouped 63
Figure 3.14 Amino Acid Alignment of Translated Genomic ABC Fragments 67
Figure 3.15 Tissue Types for Expression Studies. 74
Figure 3.16 Genomic and EST Transporter Sequences from *N. lolii* 76
Figure 3.17 Pre-optimisation for EST Primer Pairs. 79
Figure 3.18 EST Primer Pairs PCR Optimisations. 80
Figure 3.19 Nucleotide Alignment of Genomic ABCs 81
Figure 3.20 Reverse Transcriptase Tissue Expressions. 84
Figure 3.21 Evaluation of Small Primer Sets 92
Figure 3.22 Presence or Absence of Transporters in Endophyte Strains 93
Figure 3.23 Phylogeny of Endophyte Strains and Presence/Absence of Transporters 94
Figure 3.24 Phagemid Containing gABC 4e a Putative ABC Transporter Gene 101
Figure 3.25 Amino Acid Alignment of gABC 4e with the C-terminal of fungal ABCs 103
Figure 3.26 Intron Splice Sites 105
Figure 3.27 Complete gABC 4e Gene sequence 106

List of Tables

Table 2.1 Strains and Plasmids 25
Table 2.2 PCR Primers for this Research Study. 30
Table 2.3 Degenerate Primer Optimisation Mastermix 31
Table 2.4 Sequencing Primers for this Research Study 35
Table 2.5 Treatment of RNA with DNase I 37
Table 3.1 PCR Primers to MFS Alignment 1 48
Table 3.2 PCR Primers to ABC Alignment 1 53
Table 3.3 Primers of ABC Alignment 2 54
Table 3.4 BLASTX Results for the 12 Fragments 66
Table 3.5 EST putative transporters found via BLAST results 69
Table 3.6 BLASTX results of known transporters against the EST database, confirmation of the database annotation 71
Table 3.7 PCR conditions for primers used in expression studies 82
Table 3.8 Pooled BAC Screening Results 87
Table 3.9 Unpooled BAC Screening Results 88
Table 3.10 Endophyte Information 95
Table 3.11 BLASTX Results of Insert gABC 4e's Predicted ORFs 98
Table 3.12 Predicted Open Reading Frames 100

Abbreviations

ABC:	ATP-binding cassette
BLAST:	Basic Local Alignment Search Tool
bp:	base pair
cDNA:	copy DNA
°C:	degree celsius
DNA:	deoxyribonucleic acid
DNase:	deoxyribonuclease
dNTP:	deoxynucleotide triphosphate
EST :	Expressed sequence tag
g :	gram
IPTG:	Isopropyl-β-d-thiogalactoside
kb:	kilobase pair
L:	litre
LB:	Luria broth
M:	mole per litre
MFS:	Major Facillitator superfamily
ml:	milliliter
mM:	millimole per litre
NRPS:	Non-ribosomal peptide synthase
ORF:	Open reading frame
PCR:	Polymerase chain reaction
PKS:	Polyketide synthase
RNase:	ribonuclease
RNA:	ribonucleic acid
RT-PCR:	reverse transcriptase PCR
SDS:	sodium dodecyl sulfate
µl:	microlitre
µM:	micromole per litre
µg:	microgram
v/v:	volume per volume
w/v:	weight per volume
X-Gal:	5- bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Chapter 1 Introduction

1.1 *Lolium perenne* and *Neotyphodium lolii*, mutualistic symbionts

1.1.1 Endophytic Interactions

Neotyphodium lolii is a filamentous fungus, one of many that colonises a host plant as an endophyte (Yoder and Turgeon, 2001). Endophytic fungi are very common and are found in almost all species of plants (Schardl *et al.*, 2004). These fungi live in a symbiotic relationship with their host, so unlike most well studied fungal-plant pathogen interactions these relationships are beneficial to both organisms.

N. lolii belongs to a group of clavicipitaceous (*Clavicipitaceae,* Ascomycota) fungal endophytes that commonly interact with cool season temperate grasses of the Poöidae subfamily. The genus *Neotyphodium* is an asexual anamorph of the sexual *Epichloë* genus. The endophyte is symbiotic with perennial ryegrass (*Lolium perenne* L.) (Malinowski and Belesky, 2000). A widely distributed forage grass; perennial ryegrass is well suited to New Zealand's wet, mild temperate climate. In New Zealand it is the pasture grass most widely used in sheep and dairy farming because it is highly digestible and palatable (Malinowski and Belesky, 2000).

Neotyphodium endophytes live out their entire life cycle within the grass, never entering a sexual cycle, and are passed on by infecting seeds as they develop on the parent plant (Sullivan and Faeth, 2004). The fungus grows in the apoplast between the cells of the grass plant. The hyphae appear to be attached to the outside of the host cell walls and hyphal growth is synchronised with growth of the extending grass leaf (Christensen *et al.*, 2004).

This mutualistic relationship confers benefits to both the host and the symbiont. The endophyte gains the nutrition it requires and utilises its host's reproductive methods to efficiently propagate itself. The grass host is protected from biotic and abiotic stress when it is infected with endophyte, compared to those grasses that are endophyte free

(Faeth and Sullivan, 2003). As a protection from herbivory and as a deterrent to insect feeding, *N. lolii* produces many bioactive alkaloids (Spiering *et al.*, 2005a).

In contrast to fungal-pathogen interactions, mutualistic endophyte associations do not elicit an obvious defence response from the host plant. Such as production of toxic secondary metabolites, which are synthesised by plants as a response to wounding or predation by insects or phytopathogenic fungi (Del Sorbo *et al.*, 2000). Adverse reactions have been observed if the endophyte is placed in a grass species that is not its usual host: the death of hyphae in the intercellular matrix, a stunting of tillers and the loss of endophyte infectivity to the plant (Christensen, 1995). The fungal endophyte has adapted to live alongside its grass host. It may be that it is able to remove any exogenously toxins produced by the host via a transport mechanism, or perhaps is able to prevent the elicitation of any defence responses.

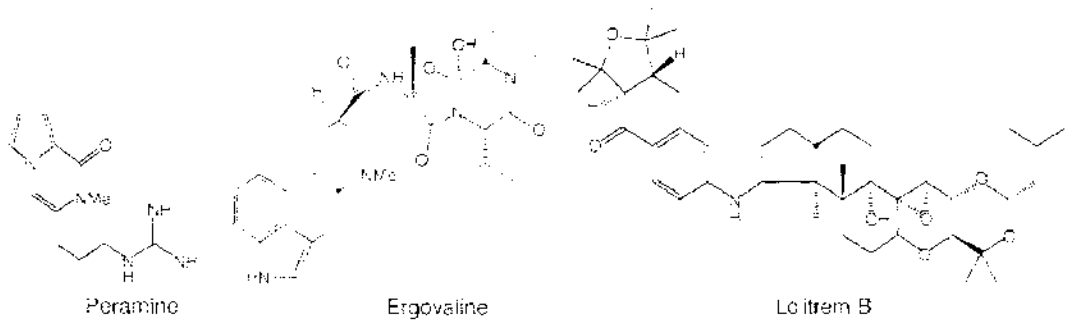
1.1.2 Alkaloid Production; Benefits to the Host

Endophyte-infected grasses are more persistent than endophyte-free grasses. They provide increased resistance to pests of the ryegrass plant, and in tall fescue confer drought tolerance under conditions of environmental stress (Malinowski and Belesky, 2000). This persistence is the reason why they are kept in agronomically important pasture. However, ingesting alkaloids such as lolitrem and peramine can poison livestock To prevent this pasture grasses have been infected with modified endophytes such as AR1 that do not produce vertebrate toxins (Malinowski and Belesky, 2000), but retain synthesis of the compounds such as peramine that enable pest resistance in ryegrass (Panaccione *et al.*, 2001).

In *Neotyphodium lolii* three major classes of alkaloids have been identified (Figure 1.1) (Panaccione *et al.*, 2003). Peramine, a pyrrolopyrazine; ergopeptines (ergovaline), and lolitrems which are indole-diterpenes (Schardl *et al.*, 2004). Peramine is an insect deterrent (Moon *et al.*, 2002), especially against Argentine stem weevil. Lolitrems are the cause of a neurological condition in livestock known as ryegrass staggers (Fletcher and Harvey, 1981) and ergot alkaloids are the cause of tall fescue toxicosis in animals ((Panaccione et al., 2001).

Endophyte toxins also protect against smaller parasites like nematodes. Kunkel (2003) looked at the effects of ergot alkaloids produced by *N. lolii* on black cutworms and their predatory nematodes. It was found that that black cutworms feeding on endophyte-infected grasses are capable of developing a resistance, by accumulation of a toxin, which protects them against their nematode predators. Loline alkaloids are broad spectrum insecticides with bioprotective effects seen against aphids. Lolines are produced by several *Neotyphodium* and *Epichloë* species but not by *N. lolii*.

Figure 1.1 Alkaloids of *N. lolii*



The three classes of biochemical alkaloids produced by *N. lolii*. Peramine and ergovaline are insect deterrents and Lolitrem B is a tremogenic mycotoxin that affects grazing mammals (Spiering *et al.*, 2005a).

1.2 Fungal Secondary Metabolites and Gene Clusters

1.2.1 Gene Clusters and the Production of Fungal Secondary Metabolites

Secondary metabolites are the low molecular weight products of plants and microbes including fungi. Production of a group of secondary metabolites is often unique to a particular species (Martin *et al.*, 2005) and the metabolites are encoded by 'dispensable' gene sets. Secondary metabolite pathways are non-essential for growth and are produced after the exponential growth phase (Martin *et al.*, 2005); or are required for growth only under sub-optimal conditions (Keller and Hohn, 1997). Secondary metabolites include compounds s such as antibiotics and mycotoxins (Keller and Hohn, 1997).

In fungi it is common for two or more genes required for the biosynthesis of one secondary metabolite to be closely linked in the genome; this is known as a gene cluster. Gene clusters usually include all the genes required in a pathway for synthesising a secondary metabolite (Keller and Hohn, 1997). Genes within a cluster generally use a common transcription factor; the genes are often situated less than 2 kb from one another (Walton, 2000). Secondary metabolites are commonly communication signals (antibiotics, mycotoxins, plant growth factors, and fungal elicitors) between the producing organism and its surrounding organisms/environment. The clusters therefore frequently contain antibiotic and toxin resistance genes or genes required to secrete metabolites as well as those that encode biosynthetic enzymes and regulatory proteins (Martin *et al.*, 2005).

The multidrug resistance (MDR) transporters of secondary metabolite pumps are of commercial interest because they secrete antibiotics such as penicillin and cephalosporin (Martin *et al.*, 2005). Genes for the biosynthesis of these industrial antibiotics (penicillin, cephalosporin, tricothecene) have been found clustered in the organisms that secrete them (Keller and Hohn, 1997). The Major facilitator superfamily (MFS) of transporters are more commonly associated with toxin biosynthetic gene clusters producing secondary metabolites than are ATP-binding cassette (ABC) transporters (Callahan *et al.*, 1999; Pitkin *et al.*, 1996). In filamentous fungi *TRI12*

(*Fusarium sporotrichioides*) and *TOXA* (*Cochliobolus carbonum*) are two secreted mycotoxins that have a MFS genes associated with their gene clusters. However some instances of putative ABC genes in secondary metabolite clusters have also been seen (Martin *et al.*, 2005).

1.2.2 Gene Clusters in Endophyte Genomes

The alkaloid compounds produced by the *Epichloë*/*Neotyphodium* genera of fungal endophytes have been well studied because of the damage they can cause to livestock and the resulting commercial losses. Recently the biosynthetic pathways that produce these metabolites have been studied for a better understanding of the metabolites' precursors, initiators and intermediates.

1.2.2.1 Lolines

Loline alkaloids which are produced in several *Neotyphodium* and *Epichloë* species, but not *N. lolii*, are broad spectrum insecticides. In *Neotyphodium uncinatum*, an endophyte that infects tall fescue, two genes *lolC* and *lolA* have patterns of expression that correlate with loline alkaloid production (Spiering *et al.*, 2002). In total the gene cluster was found to have nine genes associated with loline biosynthesis (Spiering *et al.*, 2005b), although none of these were identified as a toxin transporter.

1.2.2.2 Lolitrems

The genomes of *N. lolii* (strain Lp19) and *Epichloë festucae* (strain F11) endophytes were searched for evidence of the lolitrem gene cluster (Young *et al.*, 2005). Lolitrems are tremorgenic mycotoxins that are responsible for the neurotropic activity known as ryegrass staggers in grazing mammals in New Zealand pastures (Clay and Schardl, 2002). Three genes of the lolitrem B gene cluster have been identified so far: *ltmG*, *ltmM*, *ltmK*. These genes are all required for metabolite biosynthesis but another 10 genes at least are predicted to be involved in the biosynthesis of lolitrem B (Young *et al.*, 2005).

Paxilline is an intermediate of several indole-diterpines including Lolitrem B. Young *et al.* (2001) cloned the biosynthetic gene cluster for paxilline out of *Penicillium paxilli* and found it contained 17 genes including one putative metabolite transporter (*paxT*) (Young *et al.*, 2001). It may be that the metabolite transporter associated with Lolitrem

B has yet to be identified; however the lolitrem molecule is highly lipophilic whilst the apoplast of the grass is aqueous, it may be that lolitrem B is not a secreted toxin. Spiering *et al.* (2005a) found the occurrence of Lolitrem B increases with leaf age and has a vertical distribution in the leaf tissue. They proposed that Lolitrem B is produced constitutively and has stability within the endophyte or plant due to the increased amount of fungus in the older plant tissues.

1.2.2.3 Ergot Alkaloids

Ergot alkaloids are derivatives of lysergic acid and three other amino acids. Originally identified in species of the fungus *Claviceps*, which produces mycotoxin-filled structures called ergots, they were responsible for ergot poisoning of people and livestock. Ergovaline, found in *N. lolii*, is derived from lysergic acid and the amino acids alanine, valine and proline. Spiering *et al.* (2005a) showed the presence of ergovaline in endophyte-infected ryegrass stem tissue (the first 5 mm of a grass tiller before the meristem) and a heterogeneous distribution in the surrounding tissues regardless of the quantity of the endophyte. In leaves, ergovaline is produced in young leaf sheaths and the level of ergovaline declines in older leaf tissue.

The ergot alkaloid gene cluster encodes the ergopeptide ergotamine rather than ergovaline as seen in *N. lolii*. This molecule contains the amino acid phenylalanine as a substitute for valine (Panaccione, 2001). The ergotamine gene cluster from *C. purpurea* has been identified with 15 genes sequenced so far (Haarmann *et al.*, 2005), but no transporter gene has yet been associated with the cluster. Nor has a transporter gene been found associated with the *N. lolii* ergovaline gene cluster (Damien Fleetwood, personal communication).

1.2.2.4 Peramine

The pyrrolopyrazine alkaloid peramine is a known insect feeding deterrent that protects grasses against larval and adult herbivory of the Argentine stem weevil (*Listronotus bonariensis*) (Clay and Schardl, 2002; Tanaka *et al.*, 2005). Peramine is present in the majority of *Epichloë* and *Neotyphodium* endophyte-infected host species (Clay and Schardl, 2002). A peramine molecule contains a lipophilic ring structure and a hydrophilic guanidinium group (Figure 1.1) and is probably derived from arginine and proline (Rowan, 1993). The ratio of peramine-endophyte is higher in the apical tissues

than the basal tissues; more peramine is seen in the leaf blades although the majority of endophyte mycelia are seen in the leaf sheath tissue (Spiering *et al.*, 2005a). The hydrophilic structure and distribution of the compound suggest peramine is translocated about the plant.

Peramine is synthesised by the non-ribosomal peptide synthase (NRPS) gene *perA* from *N. lolii* and *E. festucae* (Tanaka *et al.*, 2005). This gene did not appear to be associated in a cluster of genes required for the production and transport of this secondary metabolite. Although many of the nearby genes show conserved synteny with other fungal genomes such as *Fusarium graminearum* and *Neurospora crassa*, *perA* is only known in Clavicipitaceous endophytes of grasses (Tanaka *et al.*, 2005). Of the genes clustered near *perA* two were found to share similarity with MFS transporters, although these genes are also seen in the fungal genomes in which *perA* is not conserved.

1.3 ABC and MFS transporters

1.3.1 Identification of Transporters in Fungi

Transporters are membrane proteins that form channels responsible for membrane translocation of a variety of substrates, from ions to whole proteins (Bauer *et al.*, 1999). Functions also include the transport of endogenous and exogenous toxins (Zwiers *et al.*, 2003). Transporter pumps protect pathogenic fungi from host specific and non-host specific compounds (Stergiopoulos *et al.*, 2003). Examples of this are seen in the potato pathogen *Gibberella pulicaris* and the wheat pathogen *Mycosphaerella graminicola* (Fleissner *et al.*, 2002; Stergiopoulos *et al.*, 2003; Zwiers and De Waard, 2000; Zwiers *et al.*, 2003). The two major classes of efflux pumps are the ATP-Binding Cassette (ABC) and the Major Facilitator Superfamily (MFS) transporters.

Genome sequencing and bioinformatics have allowed for the rapid discovery of numerous transporters across all organisms and accelerated the identification of transporter genes in many commercially important plant pathogenic filamentous fungi. As ABC transporters contain highly conserved motifs, and along with MFS transporters, database searching for these proteins brings up many putative transporter genes. *Neurospora crassa* is thought to contain 39 ABC transporters and 115 MFS, *Candida albicans* has 71 MFS genes and 27 ABC and a close relative of *N. lolii* the *Magnaporthe grisea* genome contains 47 ABC and 175 MFS genes (Braun *et al.*, 2005; Yoder and Turgeon, 2001).

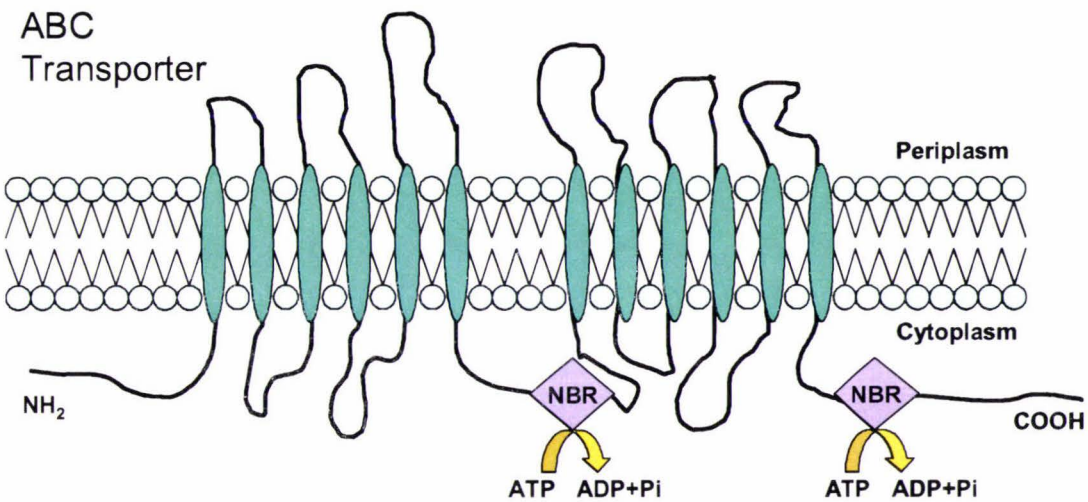
1.3.2 ABC Transporters

ABC transporters are often involved in the process of multi-drug resistance (MDR) (De Waard, 1997). They have redundancy within the genome and do not have a high degree of substrate specificity (Schoonbeek *et al.*, 2003; Zwiers *et al.*, 2003). This allows them to remove a variety of exogenous toxins (both natural and synthetic) that may be found in the environment (Andrade et al., 2000a; Hayashi et al., 2002a; Zwiers and De Waard, 2000). Drug resistance occurs when an organism has an increased number of ABC transporters that can export specific toxins from the cells (De Waard, 1997).

1.3.2.1 Structure of ABC Transporters

ABC transporters are able to bind and hydrolyse ATP molecules. The resulting energy enables solutes and macromolecules to be transported across the membrane. This is a primary active transport system (Higgins, 1992). ABC transporters generally have two regions (intracytoplasmic nuclear binding domains) that hydrolyse ATP. Two hydrophilic regions of six transmembrane domains each form a channel (Figure 1.2). In fungi the transporters are translated from a single gene that is 3900-4800 bp in size (Higgins, 1992). The nucleotide binding (NBR) domains contain motifs conserved across all ABC transporters: Walker A and Walker B regions with an ABC signature motif between them. These motifs are involved in binding of the ATP molecule.

Figure 1.2 Structure of an ABC Transporter



This transporter shows its two components each containing six transmembrane domains and a nucleotide-binding region (NBR). The NBR has the ability to hydrolyse ATP molecules into ADP and inorganic phosphate (Pi). The NBR contains the conserved motifs Walker A and B and the ABC motif. This ABC transporter shows a configuration of $(TMD_6\text{-}NBF)_2$ making it a representation of the MDR subclass.

1.3.2.2 Classes of ABC Transporters

Two main sub-classes of ABC transporters have been identified in filamentous fungi. These are multi-drug resistant (MDR) transporters and pleiotropic drug resistant (PDR) transporters. The difference between these sub-classes is structural. The arrangement of these nucleotide binding regions (NBR) and transmembrane domains (TMD) determines the class of ABC transporters. The order of domains in PDR transporters is $(NBR\text{-}TMD_6)_2$, whilst MDR transporters are $(TMD_6\text{-}NBR)_2$ as shown in Figure 1.2

(Schoonbeek *et al.*, 2003; Stergiopoulos *et al.*, 2002a). Multidrug resistance related proteins (MRPs) are a third sub-class of ABC transporter, these have the same forward orientation as MDRs $(TMD_6\text{-}NBF)_2$, but also contain additional hydrophobic N-terminal extensions (about 220 amino acids) and internal regulatory domains (Martinoia et al., 2002; van den Brule and Smart, 2002). ABCA are a fourth subclass of mammalian ABC transporters: they contain a regulatory (R) domain that sits within the membrane between the two groups of six membrane spanning domains. ABCA have not been found in yeast and only a few genes of this class have been found in filamentous fungi (Broccardo *et al.*, 1999). Half sized ABC genes are also found in genomes, the proteins are thought to dimerise to form a functional transporter channel (Stergiopoulos *et al.*, 2002b). Full sized transporters are believed to originate from the duplication of the half sized genes

1.3.2.3 Mechanism of ABC Transporters

Transporter genes are involved in a complex system to ensure the swift removal of toxic substances from within the cell (Vermeulen *et al.*, 2001). ABC transporter channels have two nucleotide-binding sites and research is being undertaken to determine how the transporters work mechanically. Biochemical studies of P-glycoprotein show both sites hydrolyse ATP molecules but the ATPase activity requires cooperative interaction between the two halves (Borges-Walmsley and Walmsley, 2001).

Schoonbeek *et al.* (2003) found that the *BcatrB* transporter from *Botrytis cinerea* recognises a range of different substrates that all contain aromatic rings. This indicates a redundancy within the toxin transport system, both for the induction of gene expression and also for transporter substrate specificity.

Binding sites of drug pumps can accommodate structurally unrelated ligands. Substrate specificity of ABC transporters appears to be determined by amino acid residues that make up the inner surface of the channel's pore. The best explanation as to how ABC transporters work is that the channel forms a large hydrophobic cavity which can hold a variety of compounds which need only to form several H-bonds to be recognised as a substrate (Tutulan-Cunita *et al.*, 2005).

1.3.2.4 Roles of ABC Transporters

Plant pathogenic fungi have to be able to protect themselves from their host plants' natural defences. Transporters may be most effective in specific spatial areas and play a role in protecting the integrity of the fungal cell membrane. Several different transporters may have substrate specificity for toxins that the plant uses for its own defence (Stergiopoulos *et al.*, 2002a). The PDR subclass of ABC transporters are involved in protection against plant defence compounds or mediate secretion of pathogenicity factors, host specific toxins or other virulence factors (van den Brule and Smart, 2002). PDR transporters are also involved in plant growth and developmental processes (Martinoia *et al.*, 2002).

ABC transporters are able to provide specific protection at certain sites within the plant. The gene *MgAtr4* encodes a PDR class transporter from the wheat pathogen *Mycosphaerella graminicola.* It is able to protect the fungus, particularly in the sub-stomatal cavities of the host (Stergiopoulos *et al.*, 2003). This transporter may act as a virulence factor by protecting the fungus against plant defence compounds.

Fleissner *et al.* (2002) studied the pathogenic fungus *Gibberella pulcaris* in its natural environment, potato tubers. It was found that MDR transporters are required to tolerate phytoalexins, which are antifungal secondary metabolites produced by the plants when they are wounded or infected by pathogens. They are also able to contribute towards pathogenicity in the tubers. ABC transporters can therefore contribute to the virulence of phytopathogenic fungi by reducing the toxic effect of phytoalexins (Fleissner *et al.*, 2002).

The sirodesmin gene cluster in *Leptosphaeria maculans* makes a class of metabolites called epipolythiodioxopiperazines (ETPs) that have antibacterial, antiviral and antifungal properties (Gardiner *et al.*, 2004). An ABC transporter gene *sirA* was found associated with this gene cluster, but as seen with *aflT*; this gene was found to be non-essential for the secretion of ETPs. However, the product of *sirA* was found to be essential for resistance of *L. maculans* to exogenous sirodesmin and gliotoxin (a related ETP) (Gardiner *et al.*, 2005). This transporter was therefore providing resistance to exogenously supplied toxins as opposed to exporting endogenous toxins.

The composition of cell walls and membranes of fungi is another area that ABC transporters and their activity may have an effect on. *MgAtr1*, *MgAtr2* and *MgAtr4* from *Mycosphaerella graminicola* were found to decrease the sensitivity of the fungus to ergosterol, the main sterol constituent of fungal membranes (Zwiers *et al.*, 2003). They suggest that ABC transporters may be involved in transporting membrane components to the outer part of the lipid bilayer to maintain the membrane, as well as to extrude any hydrophobic compounds that may disturb the membrane. Stergiopoulos *et al.* (2002a) found that the expression patterns of *MgAtr4* and *MgAtr5* are different depending on the dimorphic state the fungus is in, yeast-like or mycelium form. These two morphological stages have significant differences in the composition of their membranes which may affect the substrate specificity and activity of ABC transporters (Stergiopoulos *et al.*, 2002a).

ABC transporters play a role in maintaining membrane integrity, and some have been demonstrated to be required for virulence. *N. lolii* may have a mutualistic relationship with ryegrass but the host plant will still be producing antifungal toxins to protect against pathogens. The presence of ABC transporters assists the fungus in its natural environment by removing these toxins.

1.3.3 MFS Transporters

MFS transporters are a superfamily of polypeptides that transport a wide range of small compounds, such as toxins across chemiosmotic ion gradients (Pao *et al.*, 1998; Paulsen, 2003; Saier and Paulsen, 2001). However toxin efflux by an MFS protein is more likely to be an endogenous toxin produced by the fungi themselves (Del Sorbo *et al.*, 2000). Therefore MFS genes provide resistance to toxins similar to the ones they are able to produce.

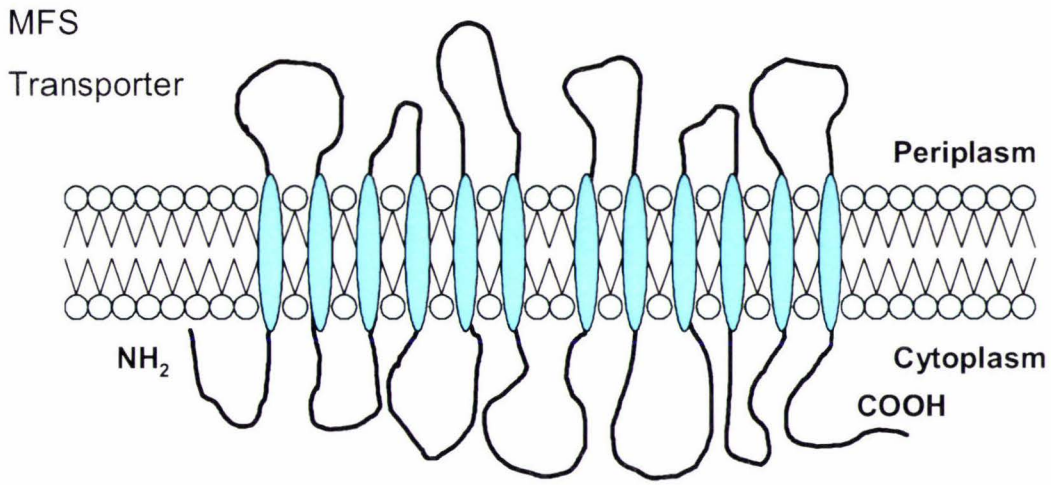
1.3.3.1 Structure of MFS Transporters

The MFS transporters are a class of transporters which use electrochemical potential gradients of charged molecules to facilitate transport of other molecular species across membranes. The MFS family transports a wide range of solutes such as sugars, peptides, drugs, organic and inorganic ions (Pao *et al.*, 1998). Protein structure is an important feature in determining the role of MFS transporters. These transporters

contain either 12 or 14 transmembrane domains (Figure 1.3) and are between 45-90kDa (Del Sorbo *et al.*, 2000; Paulsen, 2003).

The MFS transporters have been grouped into 18 different subfamilies which are identified by unique amino acid motifs. There is no one amino acid motif that links all the subfamilies together as MFS genes; rather it is the structure of 12-14 TMDs the mechanism of transport and the types of substrates transported.

Figure 1.3 Structure of a MFS Transporter



MFS transporters are composed of 12 or 14 transmembrane domains. Each subclass within the transporter family (drug transporters, ion transporters etc.) has its own amino acid motifs that identify it. There is no overall motif that identifies a MFS gene product.

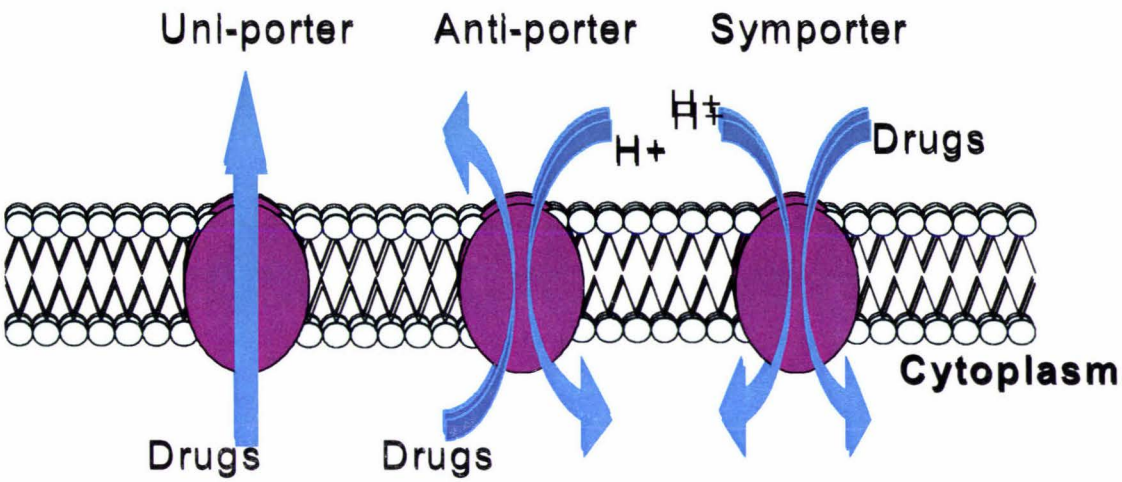
1.3.3.2 Classes of MFS Transporters

Gene clusters in fungi or bacteria that encode enzymes for secondary metabolite biosynthesis often include an MFS-like gene which has the ability to transport the final product, often a toxin, out of the cell (Upchurch *et al.*, 2002; Voss *et al.*, 2001). At least 18 different families of MFS transporter have been identified by screening peptide and translated nucleotide databases (Pao *et al.* 1998). The channels with only 12 domains are involved in the transport of sugars, peptides and ions. The 14 domain transporters are the ones that are linked with the transport of host specific toxins and mycotoxins (Stergiopoulos *et al.*, 2002b). The families are grouped mostly due to their substrate specificity although there is some conservation of sequence within these groupings.

1.3.3.3 Mechanism of MFS Transporters

The MFS transporters work as uni-porter, anti-porter or a symporters using the chemi-osmotic ion gradients to transport molecules across membranes (Pao *et al.*, 1998). They transport their bound substrate using the proton-motive force working as both influx and efflux transporters and transporting one or more substrates (Figure 1.4). Drug efflux transporters such as TetA(B) a tetracycline resistance protein from *Escheria coli* and QacA a multidrug transporter from *Staphylococcus aureus* work as drug/H+ antiporters (Ginn *et al.*, 2000).

Figure 1.4 Mechanisms of Substrate Transport by MFS



MFS transporters transport substrates across the plasma membrane as passive transporters. The movement of substrates (indicated as drugs in the diagram) is driven by chemi-osmotic energy and may also require the transport of a proton molecule (H+). Uni-porters transport the substrate alone, anti-porters move the substrate one way while transporting a proton in the other direction and symporters move both the substrate and the proton in the same direction.

1.3.3.4 Roles of MFS Transporters

A number of MFS transporters in filamentous fungi are involved in protection against a wide range of chemicals and there is significant overlap and redundancy in the molecules transported. *Bcmfs1* is a multidrug transporter gene from *Botrytis cinerea* (Hayashi *et al.*, 2002a), which transports many molecules including alkaloids, cercosporin and sterol demethylation inhibitor (DMI) fungicides. Bcmfs1 may protect *B. cinerea* against plant defence compounds during pathogenic growth, and against fungitoxic antimicrobials during saprophytic growth. A double replacement of *Bcmfs1* and another ABC gene *BcatrD* was more sensitive to DMI fungicides than a single

BcatrD gene (Hayashi *et al.*, 2002b). This shows that Bcmfs1 has a substrate overlap in its transport of natural toxins and fungicides with BcAtrD. In this set of experiments an over-expression mutant was seen to have increased sensitivity to the compound cycloheximide. An increase in sensitivity to an unrelated type of compound like this may be because MFS transporters can act as both efflux and influx transporters (Hayashi *et al.*, 2002a).

Cercospora are fungal species that produce the toxin cercosporin. They are pathogens of important crop species that cause disease symptoms by initiating oxygen-catalysed peroxidation of plant cell membranes, causing leakage and cell death. Viable hyphae were observed to be capable of reducing concentrations of cercosporin toxin from inside the fungal cell. A protein identified to be involved with resistance was the CFP (Cercosporin facilitator protein) transporter belonging to the MFS family (Callahan *et al.*, 1999). Upchurch *et al.* (2001) showed over-expression of both the *CFP* protein and transcript occurred when multiple copies of the *cfp* gene were introduced into the genome. However no change was seen in the fungus' resistance to exogenous cercosporin toxin (Upchurch *et al.*, 2002).

MFS genes are often found located close to gene clusters but they are not always involved in removing the compound from the cell. Voss *et al* (2001) identified a MFS transporter close to the gibberellin (GA) biosynthesis gene cluster in *Gibberella fujikuroi*. However the identified transporter *smtI*, contained only 12 TMDs and seems to act as an alcohol sugar transporter. Gene replacement experiments showed *smtI* was not involved in GA secretion. Voss *et al.* (2001) suggested that GA may not toxic to the fungi cells and so does not require a transporter for immediate secretion.

To establish experimentally that transporter genes found within or nearby to the gene cluster are in fact responsible for the efflux of these secondary metabolite products is a difficult task. Functional redundancy is high within these classes of genes and disruption of one putative transporter often has no effect on secondary metabolite secretion (Gardiner *et al.*, 2005). Aflatoxin is a mycotoxin produced by fungi in the *Aspergillus flavus* group of species. This compound is a known contaminant of stored feed. The biosynthetic genes for this pathway are clustered and contain *aflT*, an MFS transporter gene. Experiments determined the majority of the aflatoxin produced by this

pathway was secreted in to the media (Chang *et al.*, 2004), suggesting a transporter gene was involved. However a deletion in *aflT* did not reduce the amount of aflatoxins secreted from the mycelia. Furthermore *aflT* was not regulated by the aflatoxin pathway specific activator like the other essential genes in the cluster (Chang *et al.*, 2004).

In summary transporter proteins may be efflux transporters for toxins that have accumulated in fungal cells (Andrade et al., 2000a; De Waard, 1997; Zwiers et al., 2003). The ability of several transporters to recognise similar substrates, their redundancy, may make determination of the channel's function difficult. MFS genes are often found in gene clusters (Callahan *et al.*, 1999; Pitkin *et al.*, 1996) that encode biosynthetic pathways for secondary metabolites, particularly toxins, which allow the fungus to colonise its host plant or confer some advantage. In *N. lolii* it is expected that ABC and MFS transporters play a role in transporting alkaloids from the fungus into the plant (Callahan *et al.*, 1999; Upchurch *et al.*, 2001). There are expected to be many secondary metabolite clusters in endophytes, over and above the currently known alkaloids of peramine, ergovaline, lolitrem and loline. It seems reasonable to believe that several clusters will include transporter genes that will facilitate export of bioactive compounds into the host. Such transporters may play a role in protecting the fungi against biotic and abiotic stress as well as assisting with the uptake of essential nutrients from the apoplast.

1.4 Identifying ABC and MFS transporter genes in *Neotyphodium lolii*

1.4.1 Identification and Analysis of Transporter Genes

A high amino acid identity is seen between ABC transporters from many different filamentous fungi (Andrade et al., 2000a; Hayashi et al., 2002a; Zwiers et al., 2003). Zwiers *et al* (2000) showed that *MgAtr1* and *MgAt2* ABC transporters from *Mycosphaerella graminicola* protect against natural toxins and xenobiotics, and also have a role in pathogenesis secreting virulence factors to the host. The very close alignments of sequences such as these and other *M. graminicola* genes MgAtr4 and MgAtr5 suggest similar substrate profiles for these transporters (Stergiopoulos *et al.*, 2002a). Aligning transporter genes from fungal species closely related to *N. lolii* and known to be involved in pathogenesis or symbiosis will enable the identification of regions encoding for drug transporters, and may provide clues as to the function of these genes.

The disruption of genes of interest by homologous recombination and then study of gene function is the current way to look at toxin sensitivity in fungi (Del Sorbo et al., 2000)a; Zwiers and De Waard, 2001). However, within fungal genomes transporter genes are abundant and largely redundant (Andrade et al., 2000a; Semighini et al., 2002). Gene knockout studies by themselves have proved largely ineffective for determining toxin substrate specificity (Hayashi *et al.*, 2002a; Zwiers *et al.*, 2003). Inactivation of one ABC gene will often give no obvious phenotype as other transporters are capable of removing the same substrates. Double mutants of two similar transporter genes may not be viable as the fungus has no protection against certain toxins (Vermeulen *et al.*, 2001).

1.4.2 Identification of Transporter Genes in *N. lolii*

The EST database already contains a number of sequences that show similarity to the transporter gene families. To catalogue as many members of the endophyte transporter superfamilies as possible, the ESTs will be identified and characterised. The EST

database is a limited resource. The sequences identified in this database as similar to known ABC and MFS genes are mostly expressed sequences. They represent a subset of genes that underwent transcription during the specific conditions of the experiment. The database is expected to contain a small fraction of transporter genes within this fungal genome and although the library covers a variety of expression conditions it is unlikely to be representative of the symbiotic processes.

The genomic PCR experiment should identify a larger range of transporter-like sequences than will be seen from the EST database. Regions of conservation from known fungal ABC and MFS transporter genes will be used to design degenerate primers from *N. lolii* genomic DNA, PCR reactions should amplify a number of fragments that contain the transporter motifs but are ABC and MFS transporter genes with a variety of functions and substrates.

1.4.3 Expression of Transporter Genes

Using RT-PCR to look at the expression of any putative transporter genes identified may suggest a role in symbiosis, nutrient acquisition or toxin efflux. Increased levels of expression *in planta* compared to that seen in culture may suggest the transporter gene product has a specific role *in planta*. For example *ABC1* from *Magnaporthe grisea* is an ABC transporter gene that is essential for the pathogenicity of the fungus *(Urban et al.*, 1999). Its role is in defence against host produced anti-microbials which in turn trigger the up-regulation of the gene *in planta*.

Conversely the basal expression level of *Aspergillus nidulans* ABC transporter gene *atrC* is high compared to that of another transporter gene found in the same fungus, *atrD*. However when the fungus is treated with fungicides the expression of both these genes increases (Andrade *et al.*, 2000b). These transporters gene both protect the cell from a wide range of toxins, and their genes may share similar regulatory mechanisms but the initially high level of *atrC* expression indicates a role in ongoing metabolic functions. It may be that no up-regulation of *N. lolii* transporter genes will be observed if the gene is constitutively expressed such as those involved in the transport of nutrients.

1.4.4 Distribution of Endophyte Transporter Genes

The high redundancy of transporter genes in fungi in general and the fact that identical compounds are synthesised by many grass endophyte species and strains suggests that many of the transporter genes identified in *N. lolii* (Lp19 strain) will have counterparts gene in other endophyte strains. A distribution study across various endophytic strains will enable a presence/absence chart to be constructed. The presence of a gene in certain strains may give an indication of the role of the transporter or the types of compounds that it carries across membranes.

1.5 Aims and Objectives

This project begins the identification of transporter genes from the endophytic fungus *Neotyphodium lolii*. These transporters may have roles in protecting the endophyte from host toxins and may also be involved in transporting secondary metabolites out of the fungus and into its host plant. This project could identify those transporters that are incorporated into biosynthetic gene clusters some of which may be involved in symbiosis with the host grass.

The objective of this project is to clone fragments of the ATP-binding cassette (ABC) and major facilitator superfamily (MFS) transporter family genes from *N. lolii* and to characterise selected putative transporter genes further. Levels of expression for these genes in culture and *in planta* will be compared. A distribution study will identify the presence or absence of each gene in various endophyte strains. One transporter plus flanking DNA will be sequenced in full.

Aim 1 Identification

To amplify, sequence and clone ABC/MFS transporter genes from *N. lolii* genomic DNA using degenerate primers designed against conserved domains. To search the *N. lolii* 'in culture' database of ESTs to identify further transporters.

Aim 2 Characterization

To compare expression of transporter genes in culture and *in planta* using RT-PCR, and to study the distribution of *N. lolii* transporter genes amongst a diverse range of endophyte genotypes.

Aim 3 Analysis

To sequence at least one transporter gene from *N. lolii*, to identify its putative start site, exons and possible function. To sequence flanking DNA and identify whether the transporter is clustered with other genes of known function.

Chapter 2 Materials and Methods

2.1 Growth and Maintenance of Cultures

2.1.1 Growth of *E. coli* Cultures

E. coli cultures (Table 2.1) were grown at 37°C overnight on Luria broth (LB) agar plates or in LB broth (Appendix I), with the appropriate selective antibiotics included. The plates were then stored at 4 °C. *E.coli* strains were stored as glycerol stocks (final concentration of glycerol 30% [v/v]) at -80 °C.

2.1.2 Growth of Fungal Cultures

Neotyphodium lolii strain Lp19 (Table 2.1) was cultured on a complete medium, potato dextrose agar (PDA) (Appendix I). Mycelium was cut from the edges of the fungal colonies for sub-culturing onto fresh plates. The cultures were incubated for two weeks at 25°C.

For nucleic acid extractions of mycelia, fresh colonies (grown as described above) were chopped very finely using a scalpel blade and then transferred into 50 ml of potato dextrose broth (PDB) (Appendix I). The flasks were incubated at 25°C, shaking at 220 rpm for 10 days. Mycelia were harvested under sterile conditions with a vacuum filter containing two layers of Whatman 3 mm paper washed with sterile water.

Table 2.1 Strains and Plasmids

Strains /Plasmid	Relevant Characteristics	Source or Reference
***Escherichia coli* strains**		
XL1-Blue MRF'	Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* [F′ *proAB lacI*^{q}ZΔ*M15* Tn*10* (Tet^{r})]	(Bullock *et al.*, 1987)
SOLR	e14^{-}(McrA^{-}) Δ(*mcrCB-hsdSMR-mrr*)*171 sbcC recB recJ uvrC umuC*::Tn*5* (Kan^{r}) *lac gyrA96 relA1 thi-1 endA1* λ^{R} [F′ *proAB lacI*^{q}ZΔ*M15*] Su^{-} (nonsuppressing)	Stratagene
Top10	F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) □80*lac*ZΔM15 Δ*lac*□74 *rec*A1 *deo*R *ara*D139 Δ(*araleu*) 7697 *gal*U *gal*K *rps*L (Str^{R}) *end*A1 *nup*G	Invitrogen
Fungal Endophyte Strains		
Epichloë festucae		
AR501	Host: *Festuca arundinacea*	(Christensen *et al.* 1993)
FL1	Host: *Festuca longifolia*	(Leuchtmann *et al.* 1994)
Fp2	Host: *Festuca pratensis*	(Christensen *et al.* 1993)
RS2	Host: *Festuca arundinacea*	(Christensen *et al.* 1998)
TF15	Host: *Lolium arundinaceum*	(Christensen *et al.* 1993)
Neotyphodium lolii		
AR1	Host: *Lolium perenne*	M. Christensen (Agresearch, NZ)
AR40	Host: *Lolium perenne*	(Christensen *et al.*, 1993)
AR48	Host: *Lolium perenne*	(Christensen *et al.* 1993)
AR66	Host: *Lolium perenne*	(Christensen *et al.* 1993)
E8	Host: *Lolium perenne*	(Schardl *et al.* 1994)
Lp2	Host: *Lolium perenne*	(Christensen *et al.* 1993)
Lp19	Host: *Lolium perenne*	(Christensen *et al.* 1993)
Plasmids		
pBluescript SK(-)	Ampicillin^{R}, *LacZ'*	Stratagene
pCR 2.1-TOPO	Ampicillin^{R}, Kanamycin^{R}, *LacZ'*	Invitrogen
pECBAC1	Chloramphenicol^{R}, *LacZ'*	Friijter *et al.*, 1997

2.2 Nucleic Acid Extractions

Two methods were used for *N. lolii* DNA extractions, the Protoplasting method (http://www.aspergillus.man.ac.uk/secure/laboratory_protocols/birch.htm) and the High Molecular Weight (HMW) method (Al-Samarrai and Schmid, 2000; modified by A. Griffiths, 2001). Both protocols gave similar results and so the HMW method was followed for later extractions as it was quicker than the Protoplast method.

2.2.1 Protoplast Method

Mycelia, grown in liquid culture as described above, were filtered and washed with 1 L of sterile water and 100 ml of OM buffer (Appendix II). The mycelial mat was then scrubbed from the filter paper and added to a flask of freshly prepared Glucanex (15 mg/mL; Novozymes, Sigma-Aldrich, St Louis, Mo, USA) dissolved in OM buffer. The flasks were incubated for three hours at 30°C, with shaking at 100 rpm.

The protoplast suspension was filtered through a cloth lined sterile glass funnel into 15 ml glass corex tubes (5-10 ml of solution per tube). To extract the DNA, 5 x STC buffer (Appendix II) was poured on top of the protoplast suspension, mixed gently, and spun for 5 minutes at 1 800 x *g*. After discarding the supernatant and resuspending the protoplasts, this step was repeated once more. Finally, the supernatant was decanted off and the protoplasts were resuspended in 500 µl of STC buffer. The protoplasts were counted using a haemocytometer.

This method typically yielded $1x10^9$ protoplasts/ml, of which 500 µl of protoplasts were transferred into eppendorf tubes and centrifuged at 78 x *g* for 10 minutes. The STC supernatant was decanted and 1 ml of lysis buffer (Appendix II) added to each tube. The pellet was resuspended and vortexed for 15 seconds, then incubated at 70°C for 30 minutes and centrifuged (13 000 x *g*) for 5 minutes. The supernatant which contained the DNA was transferred into a fresh microcentrifuge tube and 300 µl of proteinase K (stock 100 mg/ml; Invitrogen, Carlsbad, CA, USA) was added and the tubes incubated overnight at 50°C.

Equal volumes of chloroform: isoamyl alcohol (24:1 [v/v]) were added to the solution, mixed by shaking and centrifuged for 10 minutes (13 000 x *g*). The aqueous layers were pooled into one tube. The DNA was precipitated by adding an equal volume of cold isopropanol (-20°C) followed by centrifugation for 30 minutes at maximum speed (13 000 x *g*). The isopropanol was decanted, the pellet air dried and the DNA resuspended in 200 µl of sterile H_2O.

The DNA was re-precipitated by adding 20 µl of 3 M sodium acetate pH 4.6 and 500 µl of 95% cold ethanol, and the tubes were spun at 13 000 x *g* for 10 minutes. The supernatant was decanted, the pellet air dried, and the DNA resuspended in 100 µl of sterile water.

2.2.2 High Molecular Weight Extraction of Endophyte Genomic DNA

Fresh fungal culture was taken from the PD broths (grown as described in 2.1.2) and washed through a Buchner funnel. Mycelia (600 mg wet weight) were ground to a fine powder using liquid nitrogen in sterile mortars and pestles. The resulting fine powder was divided between four microcentrifuge tubes (1.5 ml). The ground samples were lysed by adding 500 µl of lysis buffer (Appendix II) with 1 µl of RNase solution added to each. This was then mixed by inverting the tube until the mix became frothy. To precipitate out any polysaccharides and remaining cell debris, 165 µl of 5 M NaCl was mixed by inversion and then centrifuged at 13 000 x *g* for 12 min at 4°C.

The supernatant (~500 µl) was transferred to a new tube where 250 µl of phenol solution was added and mixed by shaking 50 times. 250 µl of chloroform/isoamyl alcohol solution was then added (24:1 [v/v]) (Appendix II), this was mixed as before then centrifuged at 13 000 x *g* for 12 minutes. The supernatant (350 µl) was transferred to a new tube and DNA precipitated with two volumes of 95% (v/v) ethanol. The solution was spun at 13 000 x *g* for 12 minutes and the ethanol drained from the tubes. To precipitate further carbohydrate, the pellet was resuspended in 250 µl of lysis buffer and 82.5 µl of 5 M NaCl added. The tube was mixed by inversion, an equal volume of chloroform (350 µl) was added, and the tube inverted again. The sample was then centrifuged at 13 000 x *g* for 10 minutes. The supernatant (about 500 µl) was then removed and transferred to a new tube. Two volumes of 95% (v/v) ethanol were added, and the DNA collected by centrifugation at 13 000 x *g* for 5-10 min. The DNA pellet

was washed by overlaying it 3-4 times with 70% (v/v) ethanol (-20°C). Finally the pellets were air dried and resuspended in 50 µl of water. Two tubes were combined to give a total volume of DNA in100 µl.

2.2.3 Isolation of Plasmid DNA from *E. coli*

Isolation of plasmid DNA from *E.coli* was carried out using a QIAGEN plasmid minikit (<20 µg plasmid DNA) (Qiagen Sciences, MA, USA) according to the manufacturers' instructions.

2.3 Molecular Techniques

2.3.1 Polymerase Chain Reaction (PCR)

The same reaction conditions were used in most experiments involving PCR and varied only for primer annealing conditions (temperature and $MgCl_2$). The amount of genomic DNA in each reaction was ~5 ng. For plasmid and BAC templates the DNA amount was 3-10 ng.

2.3.1.1 Standard PCR

PCR reactions were set up by making a master mix of all the reagents common to the reactions. PCR reaction volumes were 15 µl for 96 well plates and 0.2 ml strip tubes (Axygen, Union City, CA, USA) or 10 µl in 384 well plates. In each reaction the final concentrations of reagents were 1 x PCR buffer, 1 or 2 mM $MgCl_2$, 0.2 mM dNTPs, 1 unit/per reaction of *Taq* DNA polymerase (all reagents from Invitrogen, Carlsbad, CA, USA) and 0.5 µM each of forward and reverse primer (Table 2.2) or 1.0 µM for degenerate primers. 1 µl of template DNA (final amount as described above) was added, sterile MilliQ water brought up the volume of the total reaction. Reactions were set up on ice, mixed well and placed in a preheated (95°C) thermocycler (Biorad iCycler, Biorad, Hercules, CA, USA or 2720 Thermal Cycler, Applied Biosystems, Foster City, CA, USA).

Below is the template followed for thermocycling. The annealing temperature used depended on the optimum melting temperature (tm) conditions for each primer pair.

Initial Step	95°C	3 mins.	
Denaturation	95°C	30 secs.	30x
Annealing	58°C	30 secs.	
Elongation	72°C	1 min.	
Elongation	72°C	7 min	
Cool and Hold at 12°C			

2.3.1.2 Degenerate PCR

The degeneracy of each primer was calculated to determine the chance of binding and amplifying a specific DNA sequence. To calculate degeneracy, each letter that represents more than one amino acid residue is multiplied (eg. N=4, R=2). A good primer is one that is less than 512 fold degenerate, a higher degeneracy means there is more chance of amplifying non-specific products.

For degenerate PCR techniques to be effective and amplify the correct sequence the primers first had to be optimised. Each degenerate primer pair was screened in five $MgCl_2$ concentrations (1 mM - 5 mM) and in an annealing temperature gradient ranging from 40°C-60°C. Master mixes (of 46 reactions) were made for each primer pair; as shown in Table 2.3, with a total of 10 µl per reaction. 5 µl of 3X $MgCl_2$ stock was then added to bring the reaction up to its final 1X salt concentration in a total volume of 15µl.

Table 2.2 PCR Primers for this Research Study.

Primer Name	Size (nt.)	Tm (°C)	Primer Sequence (5' – 3')	Source
M13 Forward	16	50	GTA AAA CGA CGG CCA G	Invitrogen
M13 Reverse	17	50	CAG GAA ACA GCT ATG AC	Invitrogen
MFS 1F (d)	18	52.6	CCN RAY GAY CCN MRN AAY	This Study
MFS 2R (d)	18	56.8	NGG NGC RTA NTG YTC NGG	This Study
MFS 3F (d)	18	56.8	CCN GAR CAN TAY GCN CCN	This Study
MFS 4R (d)	15	43.2	RAA NAR NCC DAT NGG	This Study
ABC 1F (d)	18	57.2	GGN GCN GGN AAR ACN ACN	This Study
ABC 2R (d)	18	51.0	YTT NCK YTG YTC NAC NBW	This Study
ABC 4F (d)	17	55.1	GGN GCN GGN AAR ACN AC	This Study
ABC 5R (d)	17	48.7	TTN CKY TGY TCN ACR TT	This Study
ABC 6F (d)	17	49.7	ACN ATH GGN GTN GAR YT	This Study
ABC 7R (d)	18	55.1	CAT CCA YTC NGC NGG RTT	This Study
MDR 1F (d)	18	49.9	GGM AAR WSH ACB ACC ATT	This Study
MDR 2R (d)	20	62.1	CGK TGY TTY TGG CCR CCR GA	This Study
MDR 3F (d)	18	56.8	CCN GAR CAN TAY GCN CCN	This Study
MDR 4R (d)	15	43.2	RAA NAR NCC DAT NGG	This Study
4a_F2	24	64.6	GGG CAC TTG AAA CGC GCC ATT GGA	This Study
4a_R2	22	61.9	GCC CGA GCT CCT CAT GTT CCT C	This Study
4c_F2	21	62.2	CAT CGC ATT TTC GTG CCC CGC	This Study
4c_R2	26	57.9	GGC CAA ACC GAA GCT TCT ACT ATT TC	This Study
4g_F2	26	59.5	CAG GGC AAG TGA GTC CAC TAT TTT CG	This Study
4g_R2	23	68.8	TGT GGA GCT GGC CGC CAA ACC CC	This Study
GA3a_4eF	24	61.5	CGA CAC GCC TTC CTC GAT TTC CTC	This Study
GA3a_4eR	22	64.3	GCG CCA ATG TCG CCG TTG TTG G	This Study
ABC D5A_F	20	55.4	CGT TGT ATG AGG CTC GTG AG	This Study
ABC D5A_R	20	56.6	TTC GGA AGT CTC GAA CCC AG	This Study
ABC FBE_F	20	59.5	GCG CCA ATG TCG CCT TTG TT	This Study
ABC FBE_R	24	59.2	CGA CAC GCA TTC TCG ATT TCC TCA	This Study
MFS E54_F	20	56.8	TCG GAT TCA CTG GGC AAC TT	This Study
MFS E54_R	20	57.3	AGT AGA TGC TCT GCA AGC CG	This Study
MFS E94_F	20	55.3	TTC GCA GAA GAC TTT CTC GC	This Study
MFS E94_R	20	54.9	AGG TCA TAA TGG ACG GCA TG	This Study
gABC 4c F1	22	57.8	GTT GTC GCC AAT GTT GCC AAA G	This Study
gABC 4c R1	21	57.7	CCA TTC ATC AGC CCA GCG CTA T	This Study
gABC 4c F2	22	59.8	ATA GCG CTG GGC TGA TGA ATG	This Study
gABC 4c R2	22	57.3	CGA TCT GTG GTC ACA TTG	This Study
gABC 4e F1	22	58.9	CTT TGC CAC GCA TCA CGG TAA G	This Study
gABC 4e R1	19	59.7	GCG CTG CCA CTC AGA CCA GTT A	This Study
gABC 4e F2	22	61.7	CGT GCG GCTGGT CAC ATT G	This Study
gABC 4e R2	18	61.5	GCT GCG GCC GAG TCC AAA	This Study
ABCd5a F1	22	59.7	CCG CGT CTC TCC GTT CAC ATA C	This Study
ABCd5a R1	22	59.6	CGAGAGAGGCGCTTGCCATTAT	This Study
ABCd5a F2	20	60.5	GGCAGGCGGCTGGAACTTTA	This Study
ABCd5a R2	20	58.6	TGGCGGCAGCAATGTTGAAA	This Study
MFSe54 F1	22	61.1	ACTGGCTTTCCTCGCGCTAGTT	This Study
MFSe54 R1	21	59	CACAACAAACAACGACCGCGT	This Study
MFSe54 F2	21	58	CATCATCTCGGCCGCTTCAAA	This Study
MFSe54 R2	22	59.4	CCACCGACAGGCTTGGGAATTA	This Study
MFSe94 F1	22	60.2	GGAGCACAGAAGGTCGGGTTTC	This Study
MFSe94 R1	21	58.9	GATCGGCATCCTCCAGCGATA	This Study

Table 2.2 continued

Primer Name	Size (nt.)	Tm (oC)	Primer Sequence (5' – 3')	Source
GSB041 Fungactin FP	22	60.6	AAC CCG AAG TCC AAC CGT GAG A	Shalome Bassett
GSB042 Fungactin RP	24	56.9	CAT GAC AGA GTT GAA GGT GGT GAT	Shalome Bassett

Table 2.3 Degenerate Primer Optimisation Mastermix

Stock solution	(µL) in 46X reaction mix	Final Concentration (1X)
10x PCR Buffer	69.0	1X
25 mM dNTPs	5.6	0.2 mM
Primers F/R (10 µM)	6.9	1.0 mM
Taq. polymerase (5 u/µL)	11.5	1.25 U per reaction
Lp19 gDNA (100 ng/ul)	2.3	5 ng per reaction
sH_20	233.6	5 µL less per reaction to account for $MgCl_2$
$MgCl_2$ (3-15 µM)	5 µL to each	1, 2, 3, 4 or 5 µM

The thermocycling conditions used for degenerate PCR are shown below; the annealing temperatures were °C: 40, 41.4, 43.8, 47.3, 52.5, 56.2, 58.6, 60.

Initial Step	95°C	3 mins.	
Denaturation	95°C	30 secs.	┐
Annealing	40-60°C	30	secs.
30x			┘
Elongation	60°C	2 min.	
Elongation	60°C	10 min	
Cool and Hold at 12°C			

10 µl of each reaction was then run on a 1.5% (w/v) agarose gel at 70 volts for 50 minutes. The lane with the clearest band at the expected fragment size (~300 bp for all degenerate primer pairs) indicated the conditions that would give the optimum PCR conditions for later experiments.

2.3.2 Gel Electrophoresis

DNA fragments were separated by size using gel electrophoresis. All gels used for DNA and cDNA analysis were 1-2% (w/v) agarose in 1 x TAE buffer run at between 60-100 V. Samples were run against 10 µl of 1 kb plus DNA ladder (Invitrogen) and stained in 0.5 mg/ml of ethidium bromide for 20 minutes after they had been run to enable visualisation in UV light on a Biorad Gel documentation system.

2.3.3 Gel Extraction

Extraction of DNA from agarose gels was achieved using the QIAquick gel extraction kit (Qiagen) according to the manufacturer's instructions. The DNA was eluted from the column into 30 µl of water, and an aliquot (5 µl) checked for purity on an agarose gel as described above.

2.3.4 Nucleic Acid Concentration

DNA and RNA were quantified using a NanoDrop ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, DE, USA). For DNA and RNA, 1-2 µl was measured for absorbance at 260 and 280 nm. Nucleic acid concentration was calculated from the absorbance at 260 nm assuming that an absorbance of 1.0 is equivalent to 50 ng/µl for DNA and 40 ng/µl for RNA.

2.3.5 Restriction Digests

DNA was digested at 37°C for 2 hours in a 50 µl volume containing 2 µg of BAC or genomic DNA, 2 µl of enzyme (20u) and 5 µl of the appropriate buffer stock (10X). A further 2 µl (20u) of enzyme was then added and the reaction left to incubate (37°C) and cut to completion overnight.

2.3.6 Ligations

All cloning done in this project was TOPO- TA cloned using the pCR 2.1-TOPO vector from Invitrogen (Table 2.1, Appendix III). The topoisomerase enzyme is covalently bound to the cut vector; it requires only the fragment of interest to be added for ligation to occur. *Taq* polymerase adds a single deoxyadenosine (A) residue to the 3' end of the products it amplifies, this bonds with an overhanging deoxythymidine (T) at the 3' end of the linearised vector and the PCR product is ligated by topoisomerase.

The reagent volumes used were half those recommended in the TOPO cloning instruction manual (Invitrogen) and were as follows: 2 µl of PCR product (gel extracted fragment); 0.5 µl of diluted salt solution; 0.5 µl of vector (TOPO 2.1) with a total reaction volume of 3 µl. The reactions were mixed and incubated at 25°C. In the case of the degenerate fragments the ligation was left overnight to increase the ligation efficiency and therefore increase the potential diversity of fragments cloned from a mixed population, otherwise the mix was left for at least 2 hours before transformation.

2.3.7 Transformations

All 3 µl of the topoisomerase ligation reaction described above was added to an aliquot of electrocompetent cells for transformation into *E. coli* (TOP10, Invitrogen). See Table 2.1 for details of strains used in this study. The solution was mixed into the cells by swirling with the pipette. The cell mixture was gently added to a chilled cuvette and placed into the electroporator. The cells were pulsed (shocked) using a BioRad gene pulser at 25 µF, 2.5 kV and 200 ohms, then 250 µl of room temperature SOC buffer (Appendix II) was mixed in by pipetting. The cells were transferred to a new tube and incubated at 37°C for 30 min. The cells were then spread onto LB plates containing the antibiotic Kanamycin sulphate (50 µg/ml, Invitrogen) for selection and 5-bromo-4-chloro-3-indolyl-b-D-galactoside (X-Gal) (60 µg/ml) for blue/white selection. TOP10 cells do not require the addition of IPTG to obtain blue/white selection. The product is expressed from the TOPO TA cloning vector in the absence of IPTG, which is an inducer of the Lac operon. Blue/white color selection can be achieved with only X-Gal added to the plates because there is no Lac promoter that needs to be inactivated. Ligation of the fragment disrupts the *lacZ* gene and prevents β-galactosidase expression when it is inserted into the vector. Different amounts of the cell mixture were spread on each plate to get an optimal number of transformants and the plates incubated overnight at 37°C.

2.3.8 Sequencing

Sequencing reactions were carried out by myself, using an ABI 3100 Genetic analyser and the reaction chemistry was BigDye Terminator v 3.1 (AppliedBiosystems, FosterCity, CA, U.S.A.). The sequencing reactions were all run with POP 7 polymer (Applied Biosystems) through a 22 cm capillary array, which had been adapted for fast

high throughput sequencing runs for short products. However the final primer walking sequencing reactions were also run under these parameters and still gave 500-800 bp of good sequence.

2.3.9 Preparation of Sequencing Template DNA

Plasmid DNA (200ng), extracted using the Qiagen extraction method, was used directly in sequencing reactions. PCR products, sequenced with one of the primers shown in Table 2.2, were first cleaned up before the sequencing procedure was followed. After checking the PCR products on an agarose gel, the remaining 8 µl of the PCR reaction was precipitated by adding an equal volume of isopropanol (8 µl) to the remaining sample, incubated at room temperature for 20 minutes and the DNA collected at 8 800 x g for 30 minutes. The supernatant was decanted immediately and the pellet was washed in cold 70% (v/v) ethanol and spun at 8 800 x g for another 20 minutes. The ethanol was then drained off and the pellet left to dry. The DNA was resuspended in 8 µl of water and stored overnight at 4°C. The sequencing reactions used 1 µl (200 ng) of PCR amplified product as the template.

The cycle sequencing reactions were set up as follows:

	1x
5X ABI buffer	1.75 µl
Primer(M13 R.10uM)	0.5 µl
Big Dye	0.5 µl
Water	6.25 µl
PCR product	1 µl
Total	10 µl

For more than one reaction a mix was set up with 1 µl (200 ng) of PCR template added to 9 µl of the reaction mix. Primers used in sequencing reactions are shown in table 2.4

The conditions for thermocycling were

Initial Step	95°C 4 mins.	
Denaturation	95°C 20 secs.	40x
Annealing	50°C 5 secs.	
Elongation	60°C 1 min.	
Elongation	60°C 10 min	
Cool and Hold at 12°C		

Table 2.4 Sequencing Primers for this Research Study

Primer Name	Size (nt.)	Tm (°C)	Primer Sequence (5' – 3')	Source
ABCend_F3a	23	56.5	GGT CCC TGT TGC AGT TGA GTT GC	This Study
ABCend_F3b	23	56.5	GCA ACT CAA CTG CAA CAG GGA CC	This Study
4eABC_f2	22	63.2	CCA ACA ACG GCG ACA TTG GCG T	This Study
4e_F2b	22	63.2	CCC ACT AAT GAG GAA ATC GAG A	This Study
4eABC_r2	25	57.9	GCC TTC CGT GAG TAC TCA AAT TTC C	This Study
4eABC_R2b	25	57.9	GGA AAT TTG AGT ACT CAC GGA AGG C	This Study
4eABC_F3	20	64	CGC CAG CAA ACT CCG AGG CC	This Study
4eABC_F3b	20	64	GGC CTC GGA GTT TGC TGG CG	This Study
4eABC_R3	23	57.5	CGA ATA GTA CCT GAC ATG CAC CG	This Study
4eABC_R3b	22	55.4	GCC ACG TAC AGT CCA TGA TAA G	This Study
4eABC_F4a	21	61.9	CCG ATG GAC CTG AAC CCT CCC	This Study
4eABC_F4b	21	61.9	GGG AGG GTT CAG GTC CAT CGG	This Study
4eABC_R4a	22	60.5	CGG CGA CCT AAG GTT CTG TTG C	This Study
4eABC_R4b	22	60.5	GCA ACA GAA CCT TAG GTC GCC G	This Study
4eABC_F5a	22	61.5	CCT AGC CCT GCA TAC TCC GAG C	This Study
4eABC_F5b	22	61.5	GCT CGG AGT ATG CAG GGC TAG G	This Study
4eABC_R5b	23	59.5	CCG GCA TCT ACA ACT CTC CAT GG	This Study
4eABC_R5b	23	59.5	CCA TGG AGA GTT GTA GAT GCC GG	This Study
4eABC_F6a	18	59.6	GTT TAC CGC GGA TGC GGG	This Study
4eABC_F6b	18	59.6	CCC GCA TCC GCG GTA AAC	This Study
4eABC_R6a	19	61.3	GTG CCA GTT ACA GGG CCC C	This Study
4eABC_R6b	19	61.3	GGG GCC CTG TAA CTG GCA C	This Study
4eABC_F7a	20	60.1	CGT GCT CGT CTT CCC GTC TC	This Study
4eABC_F7b	20	60.1	GAG ACG GGA AGA CGA GCA CG	This Study
4eABC_F8a	20	60.8	CCA GAT AAC GCG GCG GCT AC	This Study
4eABC_F8b	20	60.8	GTA GCC GCC GCG TTA TCT GG	This Study
4eABC_F9a	22	60.5	GGC ACT CAG GAG CAT GGT TGA C	This Study
4eABC_F9b	22	60.5	GTC AAC CAT GCT CCT GAG TGC C	This Study
4eABC_F10	23	63.4	CAT GTC ATC GCT GCC GCT GCA TG	This Study
4eABC_F11	20	62.8	CAC TCG CCC ACT CCC TCT GC	This Study
ABC4e_F12	24	54.2	GGT CCC TGC GTA TAA GGG TAT TCC	This Study

M13 forward and reverse primers as well as the primers GA3a_4eF, GA3a_4eR, ABC FBE_F, ABC FBE_R primers from Table 2.2 were also used for sequencing.

The sequencing reactions were then cleaned up by adding 25 µl of a 1:1:23 mix of H_20, 3 M NaOAc and ethanol then left to precipitate at room temperature for 20 min. The samples were then centrifuged for 30 min at 8 800 x *g*, the supernatant removed, and the pellet washed with 70% (v/v) ethanol. After centrifugation for 10 min, the pellet was air dried and resuspended in 10 µl of HiDi-formamide (Applied Biosystems). The sequencing reactions were resolved on an ABI Prism 3100 Genetic Analyser, (Applied Biosystems) as described earlier in this section.

2.4 RNA Procedures

All preparations for RNA work were undertaken using RNase free plastics. All glassware was autoclaved, as were any lab prepared solutions that were to be used, including water. These were opened just prior to use. The lab surface was cleaned with a RNA protector solution RNase Zap (Ambion Inc., Austin, TX, USA), which removes RNase contamination from all surfaces. All pipettes and racks which were to be used were also treated. Disposable gloves were renewed with each handling of the tubes in the procedures and all pipette tips were RNA dedicated and opened just prior to use. All smaller RNA and cDNA preparations were undertaken in clear capped 200 µl thin walled tubes (Biorad) that were exclusively used for RNase free work

2.4.1 RNA Extraction

RNA was extracted from samples using the guanidine isothiocyanate/phenol reagent TRIzol (Invitrogen). The method recommended by the manufacturer was followed. Liquid nitrogen was used to grind 1 g of wet tissue for each sample. In the case of plant material this had been stored at -80°C, and fresh cultured mycelia was ground directly after filtering through a Buchner funnel to remove the spent culture medium. The optional isopropanol/high salt buffer (0.8 M Sodium citrate, 1.2 M NaCl) wash in the TRIzol procedure was necessary to remove additional proteoglycans and polysaccharides from these samples, particularly the plant tissues.

Several of the pellets did not adhere to the side of the tube or remained floating in the solution. Another spin was required at 10 000 x *g* for 10 min. The remaining ethanol was decanted and the pellets were air dried for 5-10 minutes. The RNA was resuspended in 200 µl of RNase-water and incubated at 60°C for 10 minutes.

2.4.2 cDNA Synthesis

Before the RNA was reverse transcribed (RT) it was first treated with DNase I (Invitrogen). 10 µg of RNA (5 µg for cDNA synthesis and 5 µg for the no RT control) was made up to a total volume of 15 µl including DNAse I and $MgSO_4$ as shown in Table 2.5.

Table 2.5 Treatment of RNA with DNase I

	RNA (10 µg) (µl)	DNase I (µl)	MgSO4 (50 mM) (µl)	Water (µl)
Uninfected leaf blade (2.5 µg/µl)	5	2	1.5	6.5
Infected leaf blade (2.3 µg/µl)	4.6	2	1.5	6.9

The RNA was incubated at 37°C for 30 minutes and then 75°C for 5 minutes.
7.5 µl (5 µg) of the DNase I treated RNA was removed to a new RNase-free tube for cDNA synthesis. cDNA was synthesised using the ThermoScript™ (Invitrogen) RT-PCR system, and the procedure was followed according to the manufacturers' instructions. The reverse transcriptase negative control contained RNA in the same volume and concentration as the cDNA reactions. The negative controls were used to confirm that no genomic DNA was present in the original RNA sample. The cDNAs and their no RT controls were then stored at -20°C.

2.4.3 Reverse Transcriptase PCR

To compare expression of individual genes in different tissue types RT PCR was used as a non-quantitative method to assay for presence or absence of mRNA transcription. RNA was extracted from *N. lolii* Lp19 growing in a liquid medium (PDA), as well as from perennial ryegrass plants both infected with endophyte, and endophyte-free. Plant samples included mature leaf blade, immature leaf blade, leaf sheath and pseudostem material. The cDNA was synthesised as described in Section 2.4.2, and the gene of interest was then amplified in a standard PCR reaction (Section 2.3.1). 1 µl of cDNA template was used in each reaction; this template was sometimes diluted depending on the strength of the bands. The PCR-amplified products from cDNA templates were first compared with their negative-RT controls to demonstrate that the amplified product was derived from cDNA and not from contaminating genomic DNA.

2.5 Library Screening

2.5.1 BAC Library Screening

The BAC library used in this study was made from genomic DNA of *N. lolii* strain Lp19 by Amplicon Express (Pullman, WA, USA). Genomic DNA was partially digested with *Mbo*I and cloned into pECBAC1 (Table 2.1). Large-scale ligations and transformations were performed and BAC clones were picked and arrayed into 384 well plates using a Biomek 2000 liquid transfer system (Beckman Coulter, Fullerton, CA, USA). Clones were grown at 37°C in LB with chloramphenicol (25 µg/ml) and glycerol stocks prepared by adding glycerol to a final concentration of 30% (w/v). The average insert size of this library is approximately 140 kb. For large scale screening, the BAC glycerols had been pooled from four 384 well plates into one 96 well plate, each well on the pooled plate corresponding to 16 separate wells on the original plates. PCR screening (Section 2.3.1) of the pooled 96-well plate was used to identify pooled wells that contained the gene of interest, and further PCR screening of the relevant 16 wells from the original 384 well BAC plate was required to narrow the search down to one positive BAC clone. BAC clones that were positive by PCR were isolated from the BAC library and streaked for single colonies on LB agar plates containing 12.5µg/ml chloroamphenicol. Individual colonies were grown overnight in LB broth containing 12.5µg/ml chloroamphenicol and rechecked by PCR as described above.

To obtain sufficient BAC DNA for further analyses, single colonies from individual BAC clones were then amplified by Multiple Displacement Amplification (MDA) using the TempliPhi™ large Construct DNA Amplification Kit (Amersham Biosciences, Little Chalfont, Buckinghamshire, UK). The supplier's protocol was followed with 2 µl of the BAC glycerol stock added to each amplification reaction. After an overnight incubation at 25 °C the samples were precipitated in 3 M NaOAc pH 4.6 and 95% (v/v) ethanol at -20°C for 2 hours and centrifuged at 8 800 x *g*. The final cleanup was done with a 70% (v/v) ethanol wash and the pellets were resuspended in 10 µl of sterile dH_2O. The final concentration of DNA in the samples was about 1 µg/µl (determined by Nanodrop spectophotometry, Section 2.3.4) and a 1:50 dilution of the samples was

amplified in a 15 µl PCR reaction to confirm the presence of the gene of interest on the BAC construct.

2.5.2 λ Library Screening

The two Lambda libraries had been created prior to this project from genomic *N. lolii* Lp19 DNA in vector Lambda ZAP® II (Stratagene, La Jolla, CA, USA). Glycerol stocks of the large (5-10 kb) and small (3-5kb) insert libraries were kept in the -80°C freezer.

These libraries were first screened by PCR (Section 2.3.1.1) to determine if the genes of interest were contained within the libraries. Most of the genes were found within both libraries and the large insert library was screened as it was more likely to contain a larger part of each gene.

Screening of the phage library meant plating out the library onto LB plates following the protocol as laid out in the Lambda ZAP® II library synthesis manual. The library was titred to allow growth of the plaques to be nearly confluent after the overnight incubation. For primary screening the library was diluted 1:10 in SM buffer and 2 µL was added to 200 µL of *E.coli* cells. For secondary and tertiary screens 2 µL of the excised plaque (in 500 µL SM buffer) was further diluted to1:100 and 1:1000.

To screen the large insert library (primary screen) following a positive PCR result, an overnight culture of *E. coli* XL-1 Blue MRF′ (Table 2.1) was prepared by inoculating a single colony into LB supplemented with 0.2% (w/v) maltose and 10 mM $MgSO_4$ followed by incubation at 30°C with shaking (200 rpm). The λ library was diluted 1:40 in SM buffer (Appendix II) and either 2 µl or 3 µl of the library was incubated with 200 µl of the overnight *E. coli* XL-1 Blue MRF′ (Table 2.1) culture for 15 min at 37°C. The phage/cell culture mix was added to 3 ml of top agarose (1.5%) equilibrated to 50°C and poured onto LB plates preheated to 37°C. The plates were incubated at 37°C for 8 hr until plaques were visible then stored at 4°C until required.

2.5.3 Plaque Lifts

Nylon filters (Hybond N+ 82mm disks; Amersham) were placed onto the plaque lawns and marked on two sides with a scalpel blade for orientation. The first lift was left for one minute to allow the DNA to adhere to the membrane, the second lift was left for

two minutes and if a third lift was taken then the membrane sat on the plate for 4 minutes.

The membranes were then placed sequentially, DNA side up, onto 3 solutions (800 µL/membrane) for DNA denaturation, neutralisation and washing. The filters were left in the denaturing solution (0.4 M NaOH, 1.5 M NaCl) for 5 minutes and in the neutralising solution (0.5 M TrisCl pH7.2, 1 M NaCl) for another 5 minutes. Baking or UV-crosslinking to fix the DNA to the membrane was not required (as the denaturation step also fixed the DNA). The membranes were transferred to 2X SSC for the wash step, dried on blotting paper and stored.

2.6 Southern Blotting with DIG-Labelled Probes

2.6.1 Labelling of DNA Probes

DNA probes were prepared by PCR amplification with a digoxigenin (DIG)-dUTP labelling kit (Roche, Mannheim, Germany) according to the manufacturer's instructions. The standard $MgCl_2$ concentration for buffers supplied with the kit was 1 mM. For those primers that were optimised to work best with slightly higher concentrations of $MgCl_2$ (2 mM), the PCR buffer was supplemented with 50 mM $MgCl_2$ stock to a final concentration of 2 mM. The probes made using these modified conditions were ABC FBE, gABC 4g and gABC 4c.

2.6.2 Hybridisation of Southern Blots and λ Plaque Lifts

Membranes were prehybridised at 42°C in 20 ml of DIG Easy Hyb solution (Roche) for at least 30 min. The DIG-labelled probe was denatured at 95°C for 3 min, snap-cooled on ice for 3 min and added to the prehybridisation buffer. All non-radioactive hybridisations were performed overnight at 42°C with gentle agitation in a hybridisation oven (Hybaid Maxi 14, Thermo Electron Corporation, Waltham, MA, USA). Membranes were washed by incubating twice, 5 min each, with 2x SSC, 0.1% (w/v) SDS at room temperature then with two 15 min washes in 0.5x SSC, 0.1% (w/v) SDS at 65°C. Following brief rinsing (1-5 min) in washing buffer (Roche), membranes were incubated for 30 min in 100 ml of Roche proprietary 1x blocking solution diluted in maleic acid (Roche) at room temperature. Membranes were incubated for a further 30

min in fresh 1x blocking solution (20 ml) containing 1µl of anti-digoxigenin Fab fragments antibodies (Roche), then washed three times (10 min each) in washing buffer.

Blots were equilibrated in 20 ml detection buffer (Roche) before 0.5-0.75 ml of the chemiluminescent substrate CDP-star (Roche), freshly diluted 1:100 in detection buffer, was applied to each Southern blot. Diluted CDP-star was equally distributed (2 ml) between 10 colony blots. Membranes were sealed in plastic bags and incubated at room temperature for 1 hr. Hybridisation signals were detected by chemiluminescence on X-ray film (Kodak BioMAX XAR) for the required exposure time and the film developed using a 100Plus™ automatic X-ray film processor (All-Pro Imaging Group, Hicksville, NY 11801).

Positive plaques were identified by aligning the X-ray film with the original plate of plaques. Using a cut-off sterile 1000 µl pipette tip, plaques were picked into 500 µl of SM buffer containing 20 µl of chloroform and incubated overnight at 4°C. The phage were diluted 1:100 and 1:1000 in SM buffer and used to prepare fresh plates for the secondary screen as described above. Positive plaques from the secondary screening were picked as before with the exception that cut-off sterile 50 µl pipette tips were used.

Excision of the pBluescript phagemid from the Lambda ZAP II vector was achieved using the ExAssist helper phage with *E. coli* strain SOLR (Table 2.1) following the manufacturers' recommendations (Stratagene). Overnight cultures of SOLR and XL-1 Blue MRF′ were grown in 50 ml of LB supplemented with 0.2% (w/v) maltose and 10 mM $MgSO_4$ and incubated at 30°C with shaking (200 rpm). The following day, the cultures were gently centrifuged (1000 x *g*) and the cell pellets resuspended in 2.5 ml of 10 mM $MgSO_4$. A mixture containing 200 µl XL-1 Blue MRF′ cells, 250 µl 2° screen phage stock and 1 µl of ExAssist helper phage was incubated at 37°C for 15 min, 3 ml of LB broth with supplements added, and the mixture incubated for 2.5-3 hr at 37°C (200 rpm). The culture was heated at 65°C for 20 min to lyse the λ phage particles and the *E. coli* cells, centrifuged for 15 min (1000 x *g*) and the supernatant containing the excised phagemids removed to a sterile tube. Excised phagemids (10 µl or 100 µl of phage supernatant) were added to 200 µl of SOLR cells, incubated at 37°C for 15 min, and 200 µl of the cell mixture spread on to LB agar supplemented with Kanamycin

50µg/ml. Plates were incubated overnight at 37°C and colonies checked by PCR for the gene of interest (Section 2.3.1).

2.6.3 Stripping Membrane-Bound Probes and Reusing Probe Solutions

The membranes were stripped of their current probes by washing in double distilled water and then two washes in DIG stripping buffer (Appendix II) for 15 minutes at 37°C. They were rinsed and stored in 2 x SSC until they were required for hybridisation with another probe.

Any probes that were to be reused were stored at -20 °C. Before reusing the probe was denatured by heating the hybridisation solution to 68 °C for 10 minutes.

2.7 Bioinformatics

2.7.1 BLAST Searches

The searches undertaken for establishing sequence identity all used the Basic Local Alignment Search Tool Algorithm (BLAST) (Altschul, 1990) and were completed on the NCBI website (http: www.ncbi.nlm.nih.gov BLAST) or on the AgResearch in-house database which is updated regularly from the NCBI site.

The BLAST algorithms used were BLASTX which translates a nucleotide sequence in all six reading frames and aligns these against a database of protein sequences. TBLASTX translates both the nucleotide sequences of interest and the chosen nucleotide database used as the comparison into all reading frames. This can improve alignment of regions that had not been identified as proteins. BLASTP was used to compare protein sequences against each other.

2.7.2 Sequence Alignments

To align nucleotide or amino acid sequences for comparison, either ClustalW from EMBL-EBI (http://www.ebi.ac.uk/clustalw/index.html) or Align X, Vector NTI (Vector NTI. suite 9; Informax) were the programs used.

2.7.3 Sequence Assembly

ABI sequencing files containing transporter sequences from Lambda library clones were imported directly into the programme ContigExpress® (Vector NTI. suite 9; Informax, Invitrogen). The sequences were trimmed of vector and poor quality sequence, and assembled into longer contiguous sequences using the defult parameters.

2.7.4 Gene Annotation

For gene annotation the gene finder programs Genezilla (Majoros, et al., 2004) (http://www.genezilla.org) and SNAP (Korf, 2004) were used to identify exons and possible gene start sites within the complete sequence of the phagemid vector insert.

InterproScan (Zdobnov, 2001) in house at AgResearch was used to locate any protein motifs and domains within the sequence that would identify any open reading frames (ORF) as putative functional proteins.

Chapter Three Results

Introduction

N. lolii is an endophytic fungus that has the ability to produce secondary metabolites such as alkaloid toxins. It is possible that these products are transported from the endophyte into the host plant and that the mechanism for metabolite mobility may be secretion by transporter genes. The endophyte also has the ability to grow unhindered within the grass; it is not affected by a host pathogenic response. Transporter genes may play a role here as well, by exporting any host-produced toxins from the endophyte cells. Transporters carry substrates other than toxins: sugars, ions and peptides. Therefore ABC and MFS transporter genes could be fundamental in maintaining the symbiotic relationship between *N. lolii* and its host *Lolium perenne*.

ABC and MFS transporter genes have been identified in all organisms studied so far. To identify them in *N. lolii* the techniques of degenerate PCR and the nesting of degenerate primers will be used to identify several transporter genes that may each be involved in carrying different substrates. In addition to this an EST library will be screened for putative transporter genes that are expressed in *N. lolii*.

The genes identified as putative ABC or MFS transporters are to be analysed by studying gene expression levels *in planta* and in culture using RT-PCR. The most interesting transporter genes are those that are up-regulated in the plant as this may indicate a role in symbiosis. On the basis of the expression study, genes of interest will be chosen for further analysis. One question to address is whether endophyte strains that differ in their natural host and in the types of secondary metabolites they produce, all contain the same identified genes. The presence or absence of transporter genes in a number of diverse strains may suggest a role in transport of a particular secondary metabolite.

The putative transporter fragments will also be used to screen BAC and lambda libraries. Transporter EST fragments will not necessarily be the same section of sequence as those amplified by degenerate PCR, and BAC screening may indicate whether ESTs localise to the same BAC clones as genomic transporter gene fragments, and are therefore linked. Complete sequencing will be undertaken for one of these genes

and the sequence analysed. The possible exons for the gene will be identified and a putative function may be assigned by comparison with other transporters using the BLAST algorithm.

This project involves identification of transporter genes that could be vital to the relationship between *N. lolii* and its host, and will indicate genes that could be researched in future studies.

3.1 Isolation ABC and MFS Transporters from Genomic DNA of *N. lolii*

3.1.1 Multidrug Transporters and the *N. lolii* Genome

When searching for transporter genes within the *N. lolii* (Lp19) genome it was hoped that several classes of both ABC and MFS genes could be recovered to gain an assessment of transporter diversity and function. The complete sequence of the *N. lolii* genome has not yet been completed therefore a degenerate PCR approach was used to recover a variety of subclasses for both MFS and ABC genes from genomic DNA. This meant that alignments of characterised transporter genes were assembled and degenerate primers designed to the conserved regions across a range of sub-families. The alignments of known ABC and MFS transporter genes were created using Clustal W and Vector NTI (Align X). The genes that formed the eventual alignments were multidrug resistance (MDR) transporters, because these genes are more commonly annotated and characterised and therefore found in the public databases and also because they are the ones of most interest in plant-fungal symbiosis as possible transporters of fungal secondary metabolites.

3.1.2 Fungal ABC and MFS Genes

The similarity between transporter gene families lies mostly in their tertiary structures; however at the amino acid level some highly conserved motifs can be identified. ABC genes have 12 trans-membrane domains that make up the channel or pore. The motifs common to all ABC transporters are the Walker A, B and ABC motifs; although these may also be found in other proteins that bind and utilise ATP. The MFS genes are poorly conserved and are broadly identified by the 12 or 14 trans-membrane domains that make up the channel. Specialised motifs for the MFS genes identify specific sub-families and their substrates. At the nucleotide level, sequence similarity is not conserved between genes as protein function is determined by amino acid motifs. Therefore degenerate primers are designed by back translation from conserved amino acids in alignments to the less conserved nucleic acid sequence.

Within conserved domains, potential primer binding sites containing amino acids with limited nucleotide redundancy were selected, for example tryptophan (TGG), or

methionine (ATG). Amino acids such as serine (TCN/AGY) or leucine (CTN/TTR) were avoided. Particular attention was paid to the 3' end of the primer binding sites where amino acids which have two set codons were selected preferentially to reduce the degeneracy at the 3' end of the primer.

3.1.3 Alignment of MFS Transporter Genes and Primer Design

To create the alignments, fungal MFS sequences were retrieved from the NCBI public database and aligned using Clustal W (Vector NTI). Few MFS transporters have been characterised in the public databases, most of the MFS genes identified by BLASTX searching with a known MFS were listed as hypothetical proteins. Several alignments were created in an attempt to identify conserved regions appropriate for degenerate primer design.

Initial alignments were created using hypothetical transporter proteins from Ascomycete fungal species such as *Aspergillus nidulans*, *Gibberella zeae*, *Schizosaccharomyces pombe* and *Candida albicans*. While small conserved regions were noted, the sequences were too diverse to contain regions with five or six conserved amino acids in a row, required to design primers. Therefore, a smaller set of MFS genes were aligned; including only multidrug resistance genes from fungal species closely related to *N. lolii*. MFS multidrug resistance proteins were used from *Gibberella pulicaris*, *Aspergillus fumigatus* (two proteins) and *Acremonium chrysogenum*. The first MFS alignment used for degenerate primer design (MFS Alignment 1) is shown in Figure 3.1.

Table 3.1 PCR Primers to MFS Alignment 1

Name	Sequence	Degeneracy	Alignment
MFS 1F	5' CCN RAY GAY CCN MRN AAY	4x2x2x2x4x2x2x4x2 = 2048	MFS 1
MFS 2R	5' NGG NGC RTA NTG YTC NGG	4x4x2x4x2x4 = 1024	MFS 1

Figure 3.1 MFS Alignment 1

Four Characterised MFS gene products aligned using Clustal W. The positions of primers MFS 1F and MFS 2R are indicated in bold. The C Motif and the H motif are enclosed in boxes. Colours indicate the level of similarity between the amino acids in the alignments. Red indicates small hydrophobic bases, blue acidic bases, brown basic bases and green is for Hydroxyl/Amine/Basic.

Unlike the ABC family, MFS genes have no motifs that are recognisable across the superfamily (Stergiopoulos et al., 2002). Individual sub-families have motifs specific to them (Figure 3.2). The WRW amino acid motif is seen in most 14 transmembrane segment MFS transporters and was the initial target for primer design. However although in these alignments it is part of the larger H motif, the other amino acids in this motif were not sufficiently well conserved for designing primers to. The primers were therefore designed to conserved regions flanking short recognisable motifs such as the H and C motifs. This ensured that the sequenced fragments would be recognised as MFS genes once they had been cloned from *N. lolii* using degenerate primers.

The position of degenerate primer pair MFS1F/2R flanking the C-motif and the WRW motif of the MFS proteins is shown in Figure 3.1.The primers are designed to genes of the 14 TMS subfamily (drug transporters) but are not designed to any MFS transporter motifs. The degeneracy of these primers was high (Table 3.1), as they were designed to maximise the number of MFS sub-families obtained.

Figure 3.2 Motifs of the MFS transporters

Motif Groups and Locations	Motif AA Conservation
All MFS Transporters	
A motif, between TMS 2 and 3 (sugar)	GXLXDrXGr-KXXIL
B motif, TMS 4 (sugar)	LiXXRXlqGXXXA
C motif, end of TMS 5 (export)	GXXXGPXiGG
14 TMS FAMILY (drug exporters)	
D1 motif, end of TMS 1	LdXtvlnvalp
E motif, TMS 17	DXXGXXL
F motif, TMS 13	IGXXXGiAvlGXl
H Motif, between TMS 5 and 6	WXwXFIINvPIG
12 TMS FAMILY (influx pumps)	
D2 motif, end of TMS 1	giglXXPvlP
G motif, end of TMS 11	GPLig

The top three motifs are found in both 12 and 14 TMS transporters, depending on the function (sugar transporter, substrate export pump). The other groups are found only in 12 TMS families or 14 TMS families. A large upper case letter indicates a conserved amino acid, lowercase indicates conservation to this amino acid or one of a similar class and an X suggests any amino acid maybe found in this position. These motifs show just how degenerate the amino acid structure of MFS transporters can be.

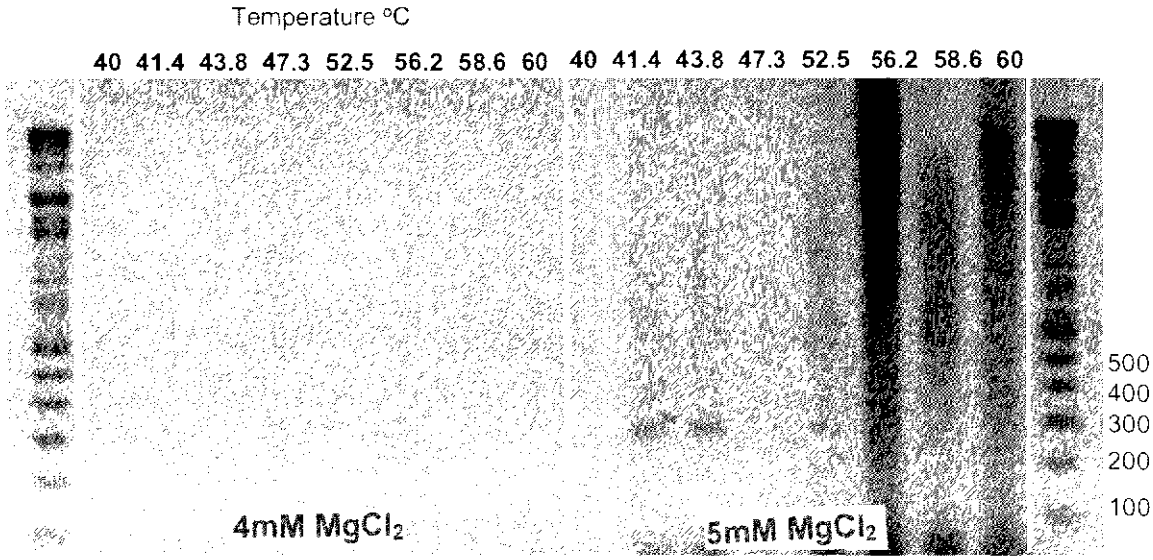
3.1.3.1 MFS PCR Optimisation

To determine conditions for generating a PCR product of the expected size (300 bp) the $MgCl_2$ concentration and annealing temperature were optimised for the primer pair. A gradient of annealing temperatures (40-60°C) was compared against a range of $MgCl_2$

concentrations (1 mM, 2 mM, 3 mM, 4 mM, 5 mM). The cycling conditions and concentration of other PCR reagents are as seen in Section 2.3.1.

The PCR products were separated by gel electrophoresis (Section 2.3.2) an example of which is shown in Figure 3.3 using PCR primers MFS 1F/2R. These primers yielded a single PCR product of 300 bp, as expected, when the PCR conditions were optimised at 43.8 °C and 5 mM $MgCl_2$.

Figure 3.3 Optimisation of $MgCl_2$ Concentration and Annealing Temperature of Degenerate Primers MFS 1F/2R

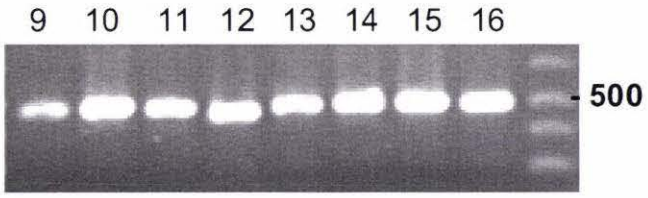


PCR products of reactions containing $MgCl_2$ concentrations of 4 mM or 5 mM. The arrow indicates the bands that were amplified at the expected size of 300 bp. The optimum annealing temperature taken as 43.8°C, the higher temperature was chosen to increase the primer stringency. All gels were run using the conditions noted in Section 2.3.2.

3.1.3.2 Cloning of Putative PCR Fragments

PCR products of the expected size (Figure 3.3) were cut from the gel, ligated into vector pCR 2.1-TOPO® (Invitrogen) and transformed into *E. coli* strain TOP10 electrocompetent cells (Section 2.3.6 and 2.3.7, Table 2.1). In all, 48 white colonies were selected and colony PCR (Section 2.3.1.1) was performed on 16 of these. Figure 3.4 shows a subset of the results. All colonies contained a DNA insert of the expected size, indicating cloning of the PCR product was successful.

Figure 3.4 PCR Confirmation of Initial Putative MFS Transformants



MFS fragments transformed into *E. coli* were tested using colony PCR with M13 F/R primers. A positive product, which contains an insert, is seen at 500 bp (100 bp vector either side of an approximately 300 bp fragment). Reactions 9-16 are shown in this figure. All white colonies picked and tested were positive for containing a fragment of the expected size.

The 16 PCR positive plasmids were extracted then sequenced (Section 2.3.8). The BLASTX algorithm was used to identify the 300 bp sequence within each positive clone. However the sequences had no recognisable or conserved pattern such as the predicted WRW motifs and were not seen to be MFS gene fragments. The BLASTX results gave no hits with an e-value less than 0.05, and no proteins were seen twice in the BLAST results of the 16 sequences. These results determined that no transporter genes had been cloned or sequenced using the degenerate primer set MFS 1F and MFS 2R.

To reduce primer degeneracy and improve the chances of cloning MFS gene fragments, alignments were constructed of MFS genes from a single sub-family. Genes of known drug transporter function and from filamentous fungal genomes closely related to *N. lolii* were selected. This second alignment included: *Aspergillus parasiticus*, potential toxin transporter; *Aspergillus fumigatus*, AflT-like major facilitator superfamily protein and multidrug resistance protein; *Gibberella pulicaris*, MFS-multidrug-resistance transporter and *Fusarium sporotrichioides,* trichothecene efflux pump. MFS Alignment 2 is shown in Appendix IV.

New primers (MFS 3F/4R) were designed to the same regions as MFS 1F and MFS 2R. By narrowing down the gene set the specificity of each primer was improved and the degeneracy reduced to 512 (forward) and 768 (reverse). The regions of similarity long enough to design primers were insufficient to allow nested primer sets. The primers MFS 3F/4R (Table 2.1) were optimised to identify the best PCR conditions but this set failed to even amplify a 300 bp fragment, despite several attempts. It was concluded

that there was insufficient conservation amongst MFS transporters to use degenerate PCR to identify MFS transporter genes from *N. lolii*.

3.1.4 Alignment of ABC Transporter Genes and Primer Design

3.1.4.1 ABC Alignment 1

As with the alignments of MFS proteins, fungal ABC transporters of known function were identified in Genbank and aligned in Clustal W (Vector NTI). The first ABC alignment contained transporter proteins previously identified from many fungal genomes; an ABC transporter from *Gibberella pulicaris*, ABC1 from *Magnaporthe grisea*, ABC transporter from *Saccharomyces cerevisiae*, PMR1 a multidrug resistance transporter from *Penicillium digitatum*, ATP-binding drug transporter from *Emericella nidulans*, CDR1 multidrug resistance gene from *Candida albicans* and an ATP-binding multidrug transporter from *Botryotinia fuckeliana* (see Appendix IV for ABC Alignment 1). These ABC proteins were all drug transporters; however the genomes ranged from close relatives to *Neotyphodium lolii* such as the genera of *Magnaporthe* and *Gibberella* to a distantly related but well studied yeast transporter.

The well conserved ATP-binding motifs (Walker A, B and ABC motifs) were identified at both the N-terminal and C-terminal ends of the protein alignments. The amino acids of the C-terminal motifs were more highly conserved, hence these were chosen for primer design. The sites picked for the degenerate primers were identical across most of the seven genes in the alignment and regions were picked that would require primers with the least degeneracy. The first alignment of fungal transporters (ABC Alignment 1) is shown in Appendix IV.

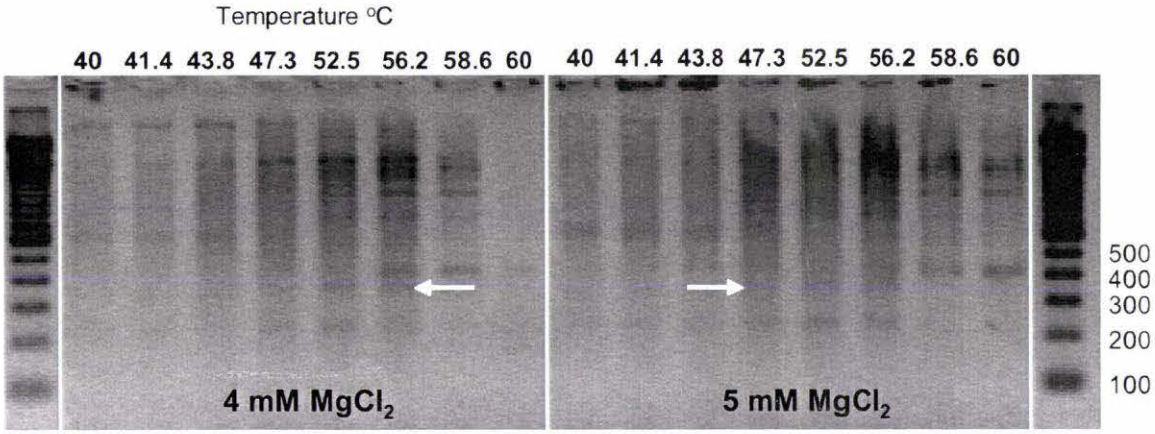
Table 3.2 PCR Primers to ABC Alignment 1

Name	Sequence (5'-3')	Degeneracy	Alignment
ABC 1F	GGN GCN GGN AAR ACN ACN	4x4x4x2x4x4= 2048	ABC 1
ABC 2R	YTT NCK YTG YTC NAC NBW	2x4x2x2x2x4x4x3x2= 6144	ABC 1

The degenerate primer pair ABC 1F and ABC 2R was designed to ABC Alignment 1 (Appendix IV) and the degeneracy of the primers was 2048 (ABC 1F) and 6144 (ABC 2R) as shown in Table 3.2. The PCR reactions were optimised for these primers as

described in Section 2.3.1.2. Figure 3.5 shows some of the results of the PCR optimisation on this primer set PCR products of the expected size, were cloned and sequenced. As described above (Section 3.1.3) the PCR products were seen on the gel as a spread of fragments (Figure 3.5), it was difficult to ensure that a single sized product had been cut from the gel. However PCR products of the expected size were sequenced to find no similarity of fragments to the expected regions of ABC transporters. The primers were found to be too degenerate to amplify the chosen sequences only.

Figure 3.5 PCR Optimisations for Primers ABC 1F/2R



Optimisation of PCR conditions at 4 mM and 5 mM $MgCl_2$ (1-3 mM $MgCl_2$ is not shown) for primers ABC 1F and ABC 2R. The different temperatures are indicated at the top of each well ranging from 40-60°C. The arrows indicate products at the expected size of 300 bp. The optimal conditions were determined to be 4mM $MgCl_2$ at the annealing temperature of 56.2°C.

3.1.4.2 ABC Alignments Two and Three

The first set of degenerate primers (ABC 1F/2R) had been designed using an alignment of seven ABC transporter genes from several different subclasses within the greater ABC family. To increase the chances of attaining transporter genes further alignments were constructed.

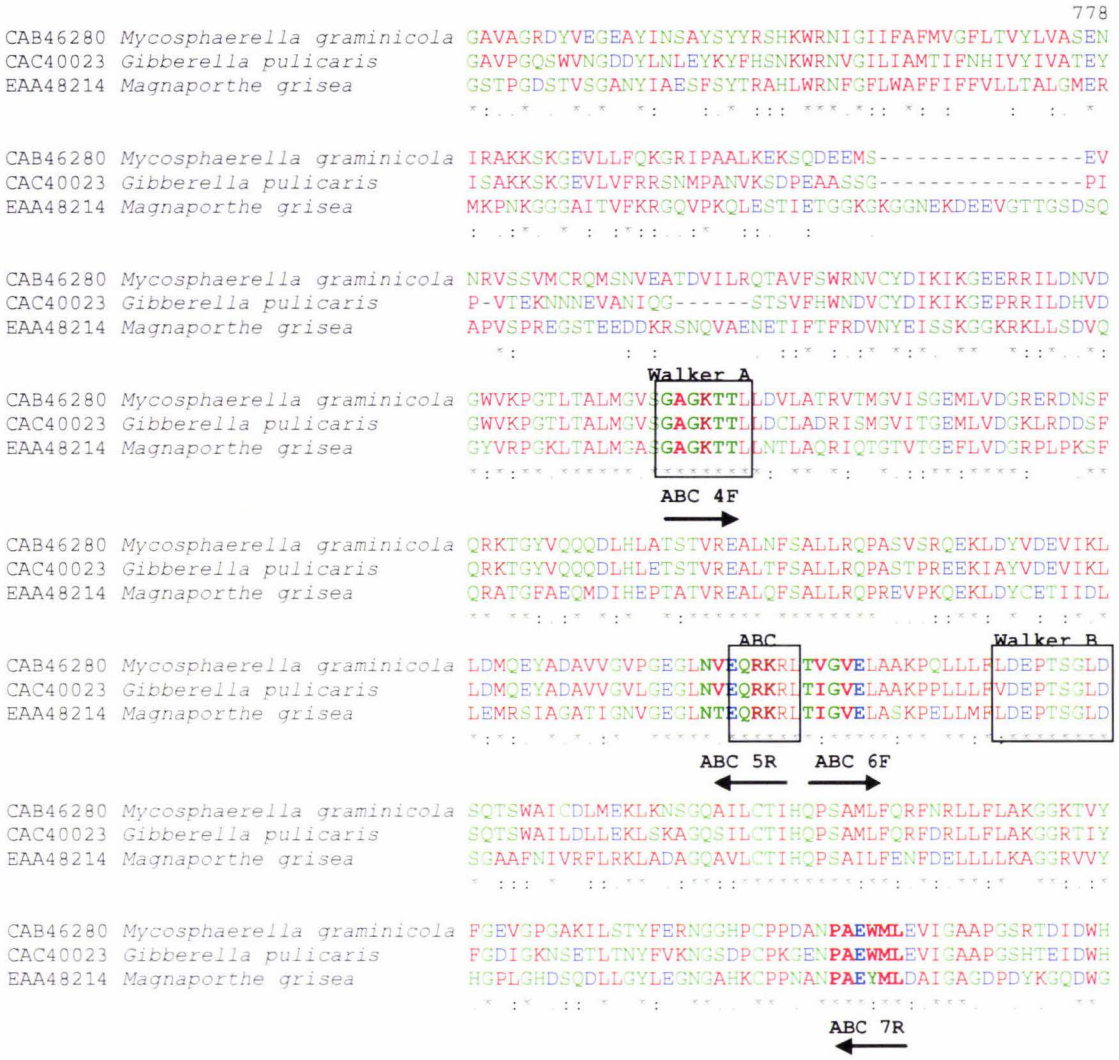
Table 3.3 Primers of ABC Alignment 2

Name	Sequence	Degeneracy	Alignment
ABC 4F	5' GGN GCN GGN AAR ACN AC	4x4x4x2x4 = 512	ABC 2
ABC 5R	5' TTN CKY TGY TCN ACR TT	4x2x2x2x4x2 = 256	ABC 2
ABC 6F	5' ACN ATH GGN GTN GAR YT	4x3x4x4x2x2 = 768	ABC 2
ABC 7R	5' CAT CCA YTC NGC NGG RTT	2x4x4x2 = 64	ABC 2

In ABC Alignment 2, regions of conservation for toxin transporters were identified in the filamentous ascomycete species *Magnaporthe grisea*, *Mycosphaerella graminicola* and *Gibberella pulicaris* (Figure 3.6).Further degenerate primers ABC 4-7 (Table 3.3) were designed to this alignment, using the principles described above.

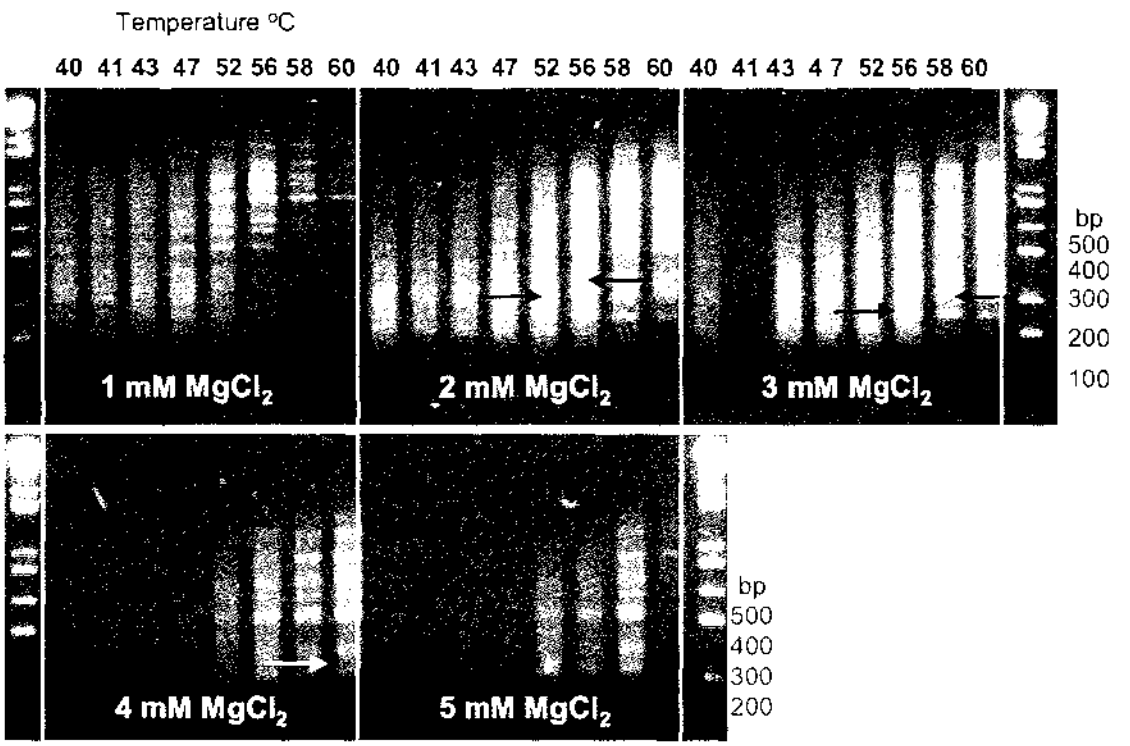
As well as ABC Alignment 2 (Figure 3.6), another set of ABC transporter proteins (ABC Alignment 3, Appendix IV) of the of multi drug resistance (MDR) sub-class were also aligned. The MDR proteins have been characterised as functional multidrug transporters belonging to the ABC class. This more specific set of primers shown in Primer Table 2.1 was included in case the set of ABC primers (ABC 4-7) from ABC Alignment 2 were still too degenerate. As before, for ABC Alignment 3 two primer pairs were designed for each transporter type to increase the chances of cloning ABC transporters. The primers were designed to conserved regions in and around the Walker A and B and ABC motifs as shown in Figure 3.6 for ABC Alignment 2 and Appendix IV for ABC Alignment 3. Optimisation of primer pairs ABC 4F/5R, with an expected PCR product size of 300 bp and MDR 1F/2R, with and expected product of 312 bp are shown in Figures 3.7 and 3.8 respectively.

Figure 3.6 ABC Alignment 2



Alignment of three ABC transporter amino acid sequences. Genbank accession numbers and species are indicated on the left. The C-terminal ATP-binding region of the gene product is shown here, with Walker A, B and ABC motifs enclosed in boxes. Primers 4F/5R and 6F/7R are indicated in bold with an arrow showing the orientation.

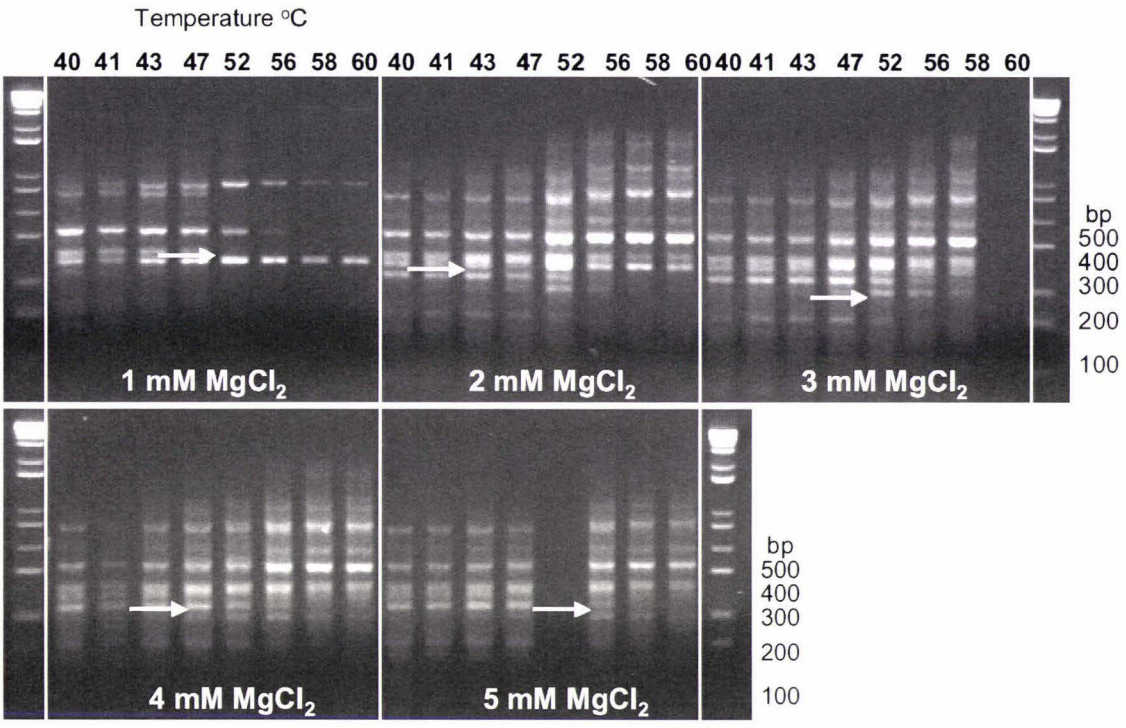
Figure 3.7 PCR Optimisation for Primers ABC 4F and ABC 5R



Optimisation of PCR conditions with primers ABC 4F and ABC 5R for $MgCl_2$ concentrations of 1-5 mM. The different temperatures are indicated at the top of the figure ranging from 40-60°C at each $MgCl_2$ concentration. The arrows indicate products at approximately the expected size of 300 bp

The products likely to contain ABC and MDR fragments were cut from the optimisation gels (Figures 3.7, 3.8). The fragments were named ABC 2-5, 2-6, 3-6, 3-7, 4-8 (Figure 3.7) and MDR 1-5, 2-3, 3-5, 4-4, 5-6 (Figure 3.8) based on the $MgCl_2$ concentration and the product size. . The ten chosen fragments (five from the ABC optimisation and five from the MDR) were purified, ligated into a pCR 2.1-TOPO® vector and transformed into *E. coli* (Section 3.1.3.2). For each of the ten fragments, between 3 and 10 white colonies were screened by colony PCR, using M13 F/R primers (Figure 3.9). A DNA product of 200 bp in size indicated an empty vector and a positive product was approximately 500 bp in size depending on the size of the fragment cut from the gel. From these ninety-six PCR positive colonies were grown overnight and used as templates for sequencing reactions as described in Section 2.3.8.

Figure 3.8 PCR Optimisation for Primers MDR 1F and MDR 2R



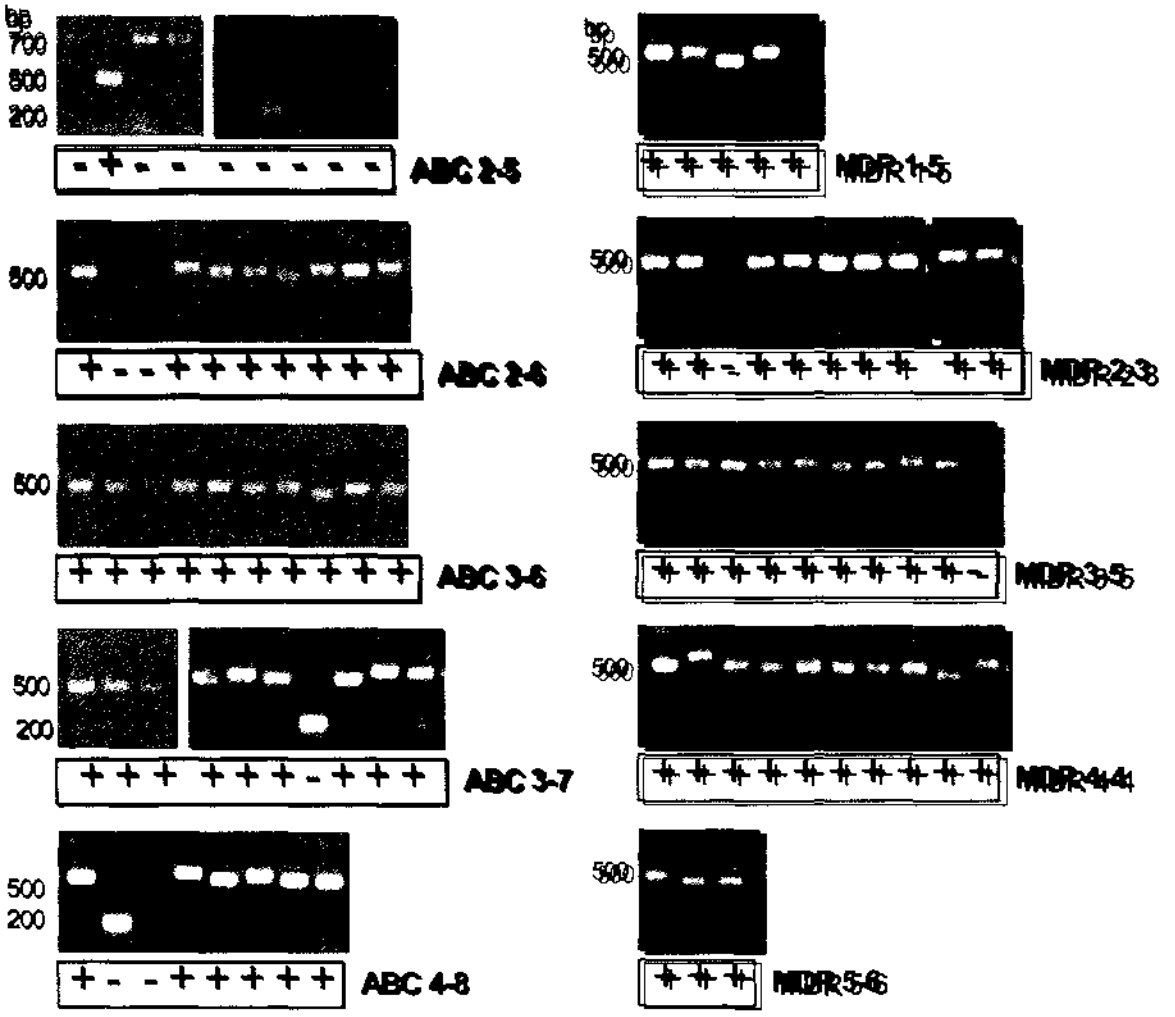
Optimisation of PCR conditions for primers MDR 1F and MDR 2R for $MgCl_2$ concentrations of 1 - 5 mM. The annealing temperatures are indicated at the top of the figure ranging from 40-60°C at each $MgCl_2$ concentration. The arrows indicate products at approximately the expected range of 312 bp.

3.1.4.3 Sequencing of Putative ABC and MDR Gene Fragments

In total 96 colonies containing ABC or MDR gene fragments were sequenced, as described in Section 2.3.8. Using an ABI 3100 Genetic Analyser (BigDye Chemistry from Applied Biosystems), running the reactions on a 22cm sequencing array, high quality sequences were obtained for 60% of the 96 clones sequenced. These were trimmed of any vector sequence and sections of poor sequencing by eye, using Chromas and Vector NTI, then compared against the NCBI database using the BLASTX algorithm. The sequences all gave different results with no significant hits to known ABCs transporters. Some results came back showing matches to: glycogen synthases (5e-43), polyketide synthases (2e-18), fungal lipotransferases and hypothetical proteins or sequences that also shared similarity with a known ABC gene product from *Gibberella pulicaris*. However some sequences also aligned with fungal MDR protein (e-0.33), canine MDR protein (e-0.45), or bacterial outer membrane protein (e-0.38). Although these low similarities could not be considered significant the results prompted improving fragment specificity using nested PCR. Since the primers were designed to

be nested should the need arise, nested PCRs were conducted as described in Section 3.1.4.4.

Figure 3.9 PCR Confirmations of ABC and MDR Transformations



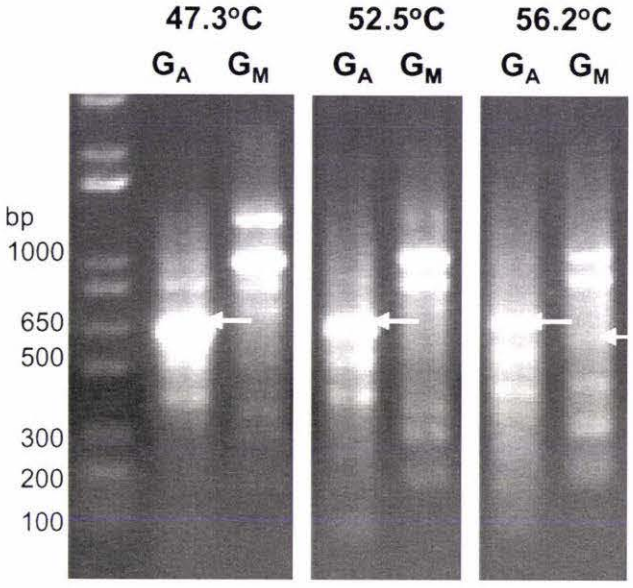
Colony PCR was used to confirm the presence of a DNA fragment ligated into vector pCR 2.1-TOPO®. The PCR DNA fragments ligated in to the vector were all about 300 bp in size, so a positive clone is approximately 500 bp and an empty vector is 200 bp.

3.1.4.4 Nested PCR Primers Improves the Identification of ABC Transporters

Nested primers were designed to the conserved ABC domains in ABC Alignments 2 and 3 to increase the specificity of the degenerate PCRs (Figure 3.6). Originally two pairs of degenerate primers were designed to each alignment, ABC 4F/5R and ABC 6F/7R (Section 3.1.4.2) for ABC Alignment 2 or MDR 1F/2R and MDR 3F/4R for ABC Alignment 3. The orientation of the primer pairs relative to each other was deliberate and enabled nesting of primers. In the first PCR primers ABC 4F/7R or MDR 1F/4R were paired and PCR amplified a product at the expected size of 600bp (Figure

3.10). Nesting the primers ABC 6F/7R and MDR 3F/4R in the second PCR, the expected product size was 300 bp. Three annealing temperatures were tested with 2 mM $MgCl_2$ concentration, as this gave the clearest PCR products when compared with 1 mM and 1.5 mM $MgCl_2$ concentrations.

Figure 3.10 Nested PCR step 1

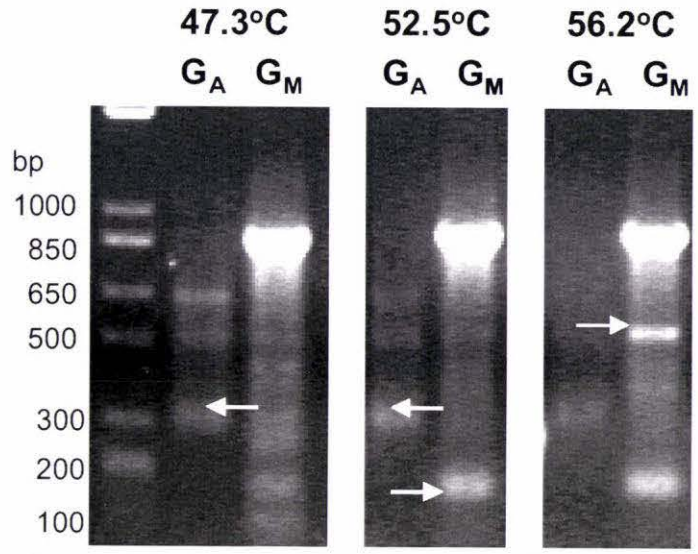


PCR of external primers (ABC 4F/7R or MDR 1F/4R) for both ABC and MDR fragments, with 2 mM $MgCl_2$. The experiment was repeated at three different annealing temperatures as indicated at the top of the figure. The samples in the wells are indicated as G_A, ABC 4F/ABC 7R (genomic DNA with ABC primers). G_M, MDR 1F/ MDR 4R (genomic DNA with MDR primers). The ladder used is 1 kb+ (Invitrogen).

In the first PCR using primers ABC 4F/7R and MDR 1F/4R, a range of DNA products of various sizes were produced suggesting non-specific binding of primers. The predominant DNA fragment for the ABC primers was approximately 600 bp as expected. These bands are shown on Figure 3.10 with arrows. A PCR product from the MDR primer pair reactions was faintly seen at 56.2 °C, this is also indicated in Figure 3.10 above. To perform the nested PCR, primers ABC 6F/7R and MDR 3F/4R were used, 1 uL of the PCR reactions from part 1 was used as the DNA template.

Four DNA amplification products indicated by arrows in Figure 3.11 were purified and cloned (Section 2.3.6). GA 300 bp at both 47 and 52°C and two MDR products GM 200 bp, GM 400 bp when no clear product was seen at 300 bp.

Figure 3.11 Nested PCR step 2



Nested PCR using an inner primer (ABC 6F or MDR 3F) and an outside primer (ABC 7R or MDR 4R). G_A, PCR part1 ABC 6F/ ABC 7R, G_M, PCR part1 MDR 3F/ MDR 4R. The red arrows indicate the fragments of interest at the expected size of 300 bp and two from the MDR reactions at 200 and 400 bp.

3.1.4.5 Sequencing Putative ABC Transporter Genes from Nested PCR Fragments

A preliminary sequencing screen of 16 clones identified that the 52°C GA 300 bp fragments (Figure 3.11) contained sequence matching ABC transporters. BLAST results for GM 400 bp gave good matches to fungal GTP cyclohydrolases, probable GTP cyclohydrolase II *Neurospora crassa* (e-value 5e-18) and GTP cyclohydrolase ii, putative *Aspergillus fumigatus* (e-value 1e-17). The MDR fragments did not share any similarity of sequence with ABC transporters, the top BLAST hits for GM 200 bp were nitrile hydratase alpha chain from uncultured bacterium (e-value 5e-07) and formate dehydrogenase beta-subunit from *Methanococcus voltae* (e-value 5e-05).

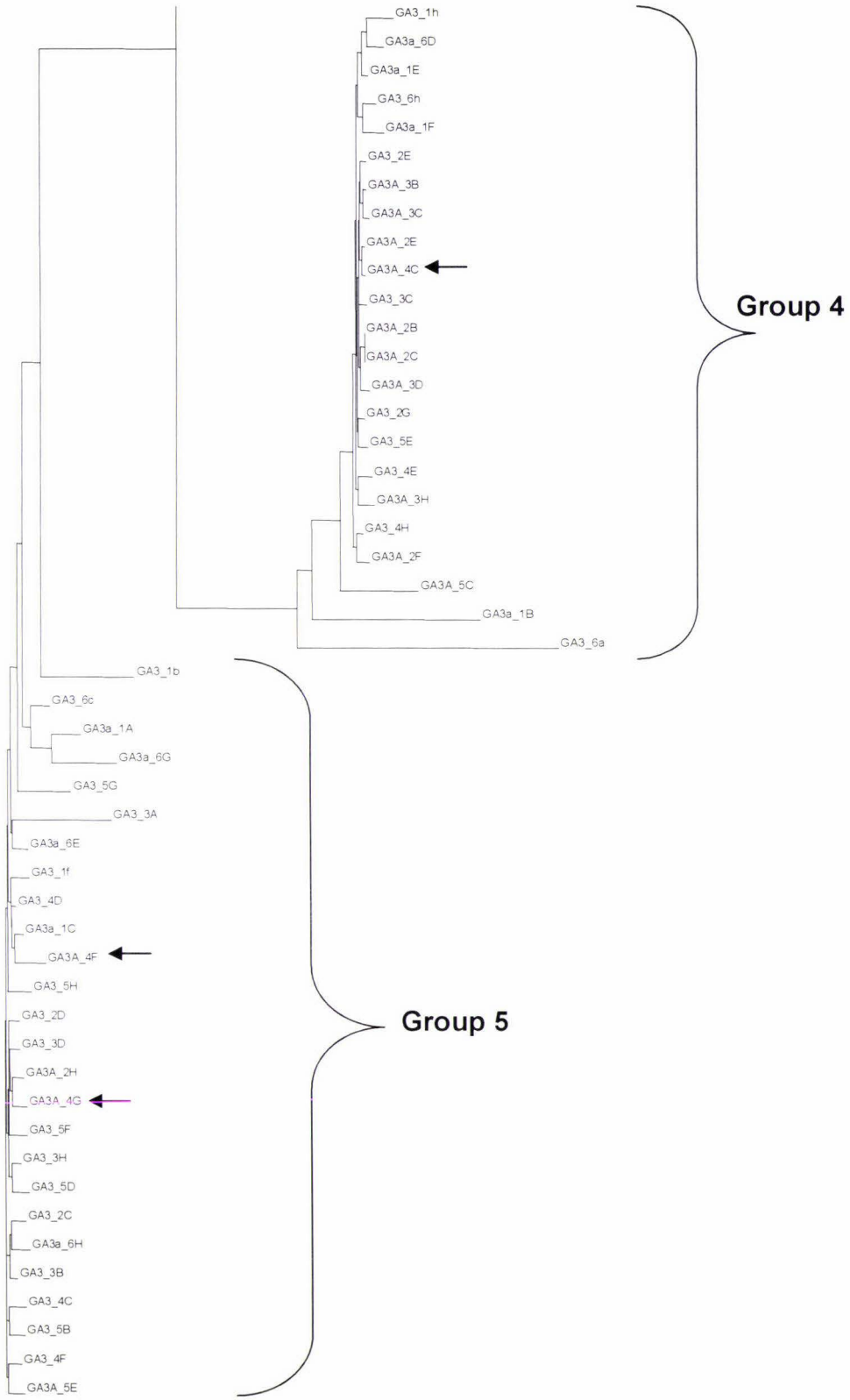
A 96 well plate of additional samples were picked off a plate of *E. coli* transformants containing GA 300 bp fragments and the colonies were grown overnight under antibiotic selection in liquid broths as described in Section 2.1.1. Plasmids were sequenced on an ABI 3100 genetic analyzer, giving 88 high quality sequences. Numerous colonies were sequenced in the hope that many of them would contain different gene fragments. Although some colonies were expected to contain the same sequence as others, additional sequencing increased the chances of finding diverse sequences.

Using the BLASTX algorithm sequences were compared against NCBI databases, such as genbank. Most *N. lolii* sequence amplified by nested primers ABC 6F and ABC 7R showed high sequence identity to regions of known ABC-transporters from fungal species such as *Mycosphaerella graminicola* (Figure 3.12), *Gibberella pulicaris* and other fungal ABC transporters. These results indicted the identification of ABC genes in genomic DNA of *N. lolii* through the nesting of degenerate primers had been successful.

Figure 3.12 BLASTX Alignment

```
>gb|AAK15314.1| ABC transporter Atr4 [Mycosphaerella graminicola]
Length = 1635

Score = 155 bits (391), Expect(2) = 3e -38

Identities = 75/90 (83%), Positives = 81/90 (90%)

Frame = -2

GA3a 1F: 322 SIXVELXAKPKLXLFLDEPTSGLDSQSSWAICAFLXXLADSGQAILCTIHQPSAILFQAF  143
            +I VEL AKPKL LFLDEPTSGLDSQS+WAICAFL  LAD+GQA+LCTIHQPSAILFQ F
 Atr4 : 1049 TIGVELAAKPKLLLFLDEPTSGLDSQSAWAICAFLRKLADAGQAVLCTIHQPSAILFQEF  1108

GA3a 1F: 142 DRLLFLAKGGKTVYFGNIGDNSRTMLDYFE 53
            DRLLFL KGG TVYFG+IG NSRT+LDYFE
 Atr4 : 1109 DRLLFLRKGGHTVYFGDIGKNSRTLLDYFE 1138

Score = 25.0 bits (53), Expect(2) = 3e -38

Identities = 9/16 (56%), Positives = 13/16 (81%)

Frame = -3
GA3a 1F: 48 SGARKCDDKEXPAEWM 1
            +GAR C ++E PAE+M
 Atr4 : 1140 NGARDCGEEENPAEYM 1155
```

Genomic sequence GA300 1F (323 bp), from nested fragment GA 300 bp located in well 1F of the 96 well plate. Aligned with ATR4 an ABC transporter from *M. graminicola* using BLASTX. The alignment matches the Walker B motif underlined (DEPTSGLD).



Grouping of the 88 sequenced GA 300 bp fragments derived from the nested PCR (Section 3.1.4.4) into clusters using Vector NTI Align X. Individual sequences within groups are replicate copies of the same sequences; the differences within are due to sequencing errors. The sequences chosen for BLASTX analysis are indicated by arrows; these can be seen in Table 3.4. Groups 1, 2, 4 and 5 are similar to ABC transporters while groups 3a and 3b have similarity with HAAAP and elongation factor gene fragments.

To sort the sequences into homologous groups, all 88 sequences derived from nested PCR with primers ABC 6F/7R were imported into Vector NTI Align X (based on Clustal W) to produce an alignment of sequences and a dendrogram showing the relatedness between sequences. Prior to alignment the 88 sequences were trimmed of all vector sequences and poorly-called bases.

The DNA fragments grouped into approximately 5 distinct clusters (Figure 3.13), with some of the more unique or poor quality sequences as outliers. Disparity between the sequences within each cluster was the result of sequencing variation. The consensus sequence for each group was chosen for analysis by BLASTX to determine putative function for that sequence. In the case of groups 2, 5 and 3b where the sequences were more diverse, two or more of the fragments were analysed using the BLASTX algorithm to ensure that the more obvious differences between these sequences were not just treated as sequencing errors.

All 12 of the selected consensus sequences (highlighted in Figure 3.13) had significant sequence identity to ABC transporter genes, or to other gene families containing ATP-binding domains as determined by BLASTX (Table 3.4). The sequences of GA3 2D (group3B), 2G (group 3A), and 3E (group 3B) all contained the Walker B motif as expected, but were closer in identity to other genes containing these motifs such as elongation factor-3 genes and hydroxy-amino acid/alcohol permease (HAAAP) genes. These proteins, like ABC transporters, utilise the energy of ATP in their function. These three different proteins all have the same recognition and binding motifs for the ATP molecule. The remaining nine sequences all had similarity to ABC transporter genes (Table 3.4). The hypothetical genes listed in the BLASTX results for GA3_3F and GA3_4A fragments (Table 3.4) do contain the Walker A, B and ABC motifs and are expected to be ABC transporter genes.

These nine sequences were realigned and shown to represent four putative ABC gene fragments. This was expected as Figure 3.10 shows five groups, one of which did not contain ABCs transporters, and some duplicate copies from each group were analysed as mentioned above. The four putative ABC transporter genes identified from the genome are GA3 4a, 4c, 4e and 4g, one from each of the groups 1, 2, 4 and 5 as shown in Figure 3.13.

Table 3.4 BLASTX Results for the 12 Fragments

Sequence ID	Top BLAST Hits (Accession number and Description)		Score	E-value
GA3_2D	AAX07692	elongation factor 3-like protein *Magnaporthe grisea*	141	6e-33
	EAA51680	hypothetical protein *Gibberella zeae*	129	3e-29
GA3_2G	YP_068880	putative HAAAP *Yersinia pseudotuberculosis*	138	6e-32
	CAC86203	putative transporter protein *Salmonella typhimurium*	169	2e-41
GA3_3E	AAX07692	elongation factor 3-like protein *Magnaporthe grisea*	126	2e-28
	EAA51680	hypothetical protein *Gibberella zeae*	112	2e-24
GA3_3F	EAA74126	hypothetical protein *Gibberella zeae*	55	6e-09
	EAA51680	hypothetical protein *Magnaporthe grisea*	46	0.0003
GA3_4A	EAA72370	hypothetical protein *Gibberella zeae*	168	5e-41
	BAA93677	BMR1 *Botryotinia fuckeliana*	140	9e-33
	BAC67162	ATP-binding cassette transporter *Magnaporthe grisea*	138	6e-32
GA3_4C	AAK15314	Atr4 *Mycosphaerella graminicola*	167	3e-42
	CAC41639	BcatrD protein *Botryotinia fuckeliana*	155	5e-39
	AAK62810	ABC1 Venturia inaequalis	167	9e-43
GA3_4E	AAR10387	P-glycoprotein 1 *Sorghum bicolor*	91	8e-18
	CAD59588	MDR-like ABC transporter *Oryza sativa*	91	1e-17
	AAL57243	ABC4 *Venturia inaequalis*	91	1e-17
GA3_4F	AAB86640	ABC1 transporter *Magnaporthe grisea*	179	2e-44
	CAC40023	ABC-transporter *Gibberella pulicaris*	167	8e-41
GA3_4G	AAB86640	ABC1 transporter *Magnaporthe grisea*	191	6e-48
	CAC40023	ABC-transporter *Gibberella pulicaris*	178	5e-44
GA3_5B	AAR10387	P-glycoprotein 1 *Sorghum bicolor*	90	2e-17
	AAL57243	ABC4 *Venturia inaequalis*	90	2e-17
GA3_5C	CAA93140	ATP-binding cassette multidrug transporter *Emericella nidulans*	148	4e-35
	AAK15314	Atr4 *Mycosphaerella graminicola*	143	1e-33
	AAK62810	ABC1 *Venturia inaequalis*	143	1e-33
GA3_6C	AAB86640	ABC1 transporter *Magnaporthe grisea*	180	1e-44
	CAC40023	ABC-transporter *Gibberella pulicaris*	167	9e-41

Table 3.4 legend BLASTX results for the 12 Fragments (page 65)

The 12 selected genomic ABC 300 fragments (GA3) from Figure 3.13 and their top two or three BLAST results with their accession numbers and description of the genes. The number and letter in the sequence ID (e.g. _2D) indicates well position in the original plate.

Figure 3.14 shows the amino acid alignments for the four putative ABC transporters GA 300 4a, 4e, 4c and 4g, which were chosen from the nested genomic fragments using alignments and the BLASTX algorithm. Sequences were translated into all six frames to detect the ABC motifs. The GA3 4a, 4c and 4g sequences code for approximately 100 amino acids including the Walker B motif. Translated GA3 4e sequence includes the Walker A motif, instead of the Walker B. Interestingly the GA3 4e fragment should not have been amplified in the nested PCR using primers ABC 6F/7R (Figure 3.6) and yet fragments containing this sequence were represented 15 times in the GA 300 transformants. GA3 4e therefore does not align with the three other sequences but is included in the figure to illustrate this point. The amino acid alignment of GA3 4e, 4a, 4c and 4g is shown in Figure 3.14.

Figure 3.14 Amino Acid Alignment of Translated Genomic ABC Fragments

```
Trans. GA3_4e (-3)     (1) ----------ANVAFVGTTGSGKS-----TIVSLIERFYDPTSGRILVDS
Trans. GA3_4a (-1)     (1) TIGVELASKPELLMFLDEPTSGLDSGAAFNIVRFLRKLADAGQAVLCTIH
Trans. GA3_4c (-1)     (1) TIGVELAAKPKLLLFLDEPTSGLDSQSSWAICAFLRKLADSGQAILCTIH
Trans. GA3_4g (-1)     (1) TIGVELAAKPPLLLFVDEPTSGLDSQTSWAILDLLEKLTKAGQAVLCTIH

Trans. GA3_4e (-3)    (36) KPVTSLRLSEYRNALVWSAK---SPPCSTAPSR-I-RLVWTKRAKHHPLM
Trans. GA3_4a (-1)    (51) QPSAVLFEHFDELLLLKSGGRVVYHGPLGEDSRTLIDYLESNGAFQVPPR
Trans. GA3_4c (-1)    (51) QPSAILFQAFDRLLFLAKGGKTVYFGNIGDNSRTMLDYFEQKRGTKMR-Q
Trans. GA3_4g (-1)    (51) QPSAMLFQRFDRLLFLAKGGKTVYFGDIGEKFQDDDELF-AKIVDSLALK

Trans. GA3_4e (-3)    (81) RKSRMRV--
Trans. GA3_4a (-1)   (101) GQPLRNG--
Trans. GA3_4c (-1)   (100) GKPRGVD--
Trans. GA3_4g (-1)   (100) AANSRXNGR
```

Clustal W alignment of the four putative ABC amino acid sequences from *N. lolii* as indicated in Table 3.4. The motifs are shown in bold, with the Walker B (DEPTSGLD) underlined in fragments GA3 4a, 4g, 4c and the Walker A (GSGKST) in bold for fragment GA3 4e. The colours indicate homology between the positions of the sequence with yellow conserved in all, blue in two or more and green for amino acids with similar properties.

3.1.5 Identification of ABC and MFS Transporters from an *N. lolii* EST Database

Transporter gene identification from the *N. lolii* genome using degenerate primers resulted in the four putative ABC gene fragments but no MFS gene fragments. The ABC genes were all similar, with BLAST results aligning the GA 300 fragments GA 4a, 4e, 4c and 4g with ABC drug transporters. In an attempt to identify different types

of transporters an EST (Expressed Sequence Tags) database was searched. The EST database is an AgResearch proprietary database developed by the Advanced Fungal Technologies team, (AgResearch Grasslands, Palmerston North New Zealand). The database contains over 3000 ESTs plus a few shotgun BAC sequences from *N. lolii* Lp19 and was made available for this experiment to be searched for transporter genes of both MFS and ABC classes.

The EST database includes the expressed cDNA sequences of *N. lolii* Lp19 endophyte cultured in both Blankenship (Blankenship, 2001) media, a minimal medium that simulates *in planta* growth, and the complete medium PDA (Appendix I) . The ESTs had been annotated *in silica*. The database was therefore able to be searched both by using sequence identity (BLAST searching) and by keyword searching of the annotations. Putative transporter genes in this library were added to the sequences identified in the genome using degenerate PCR.

The database was searched for MFS and ABC transporter sequences using two techniques. Firstly ABC and MFS genes were identified from Genbank (NCBI) using a keyword search and the corresponding proteins were used to identify transporter sequences from the database by virtue of homology. The annotated EST database was also searched using keywords, such as searching for 'ABC transporter', 'MFS transporter' and 'Drug or Toxin transporter'. It was also possible to view a table of annotations and identify any BLAST matches that might relate to transporter genes of interest. Sequences previously aligned in this project to create alignments for designing primers were also used; including both ABC and MFS genes from fungal species closely related to *N. lolii* such as *Gibberella pulicaris*, *Aspergillus fumigatus* and *Magnaporthe grisea*.

Using both these methods four putative transporter genes were identified from the EST database: two ABC genes and two MFS genes. The genes were checked again against the NCBI databases to confirm that transporters had been identified and the annotations were correct. The four EST sequences are listed in Table 3.5 along with their transporter type and top BLAST hits. Table 3.5 notes the putative transporters as belonging to the ABC or MFS classes and percentage amino acid identity is shown (% ID) as it relates across the length of the EST to the compared protein.

Table 3.5 EST putative transporters found via BLAST results

EST ID	Transporter Type	BLASTP Results	% ID	e-value
CS2400000 1FF**D5A**	*ABC*	AB164460 *Penicillium digitatum* PMR4 ATP binding cassette transporter	*58*	*e-100*
		AF032443 *Magnaporthe grisea* ABC1 transporter	*72*	*2e-99*
CS2400000 1FF**FBE**	*ABC*	FNU62929 *Filobasidiella neoformans* mdr1	55	1e-62
		SFU62933 *Aspergillus fumigatus* mdr1	49	1e-62
CS2400000 1FF**E54**	*MFS*	POL243458 *Penicillium olsonii* mdr gene	56	3e-07
		GPU132188 *Gibberella pulicaris* mfs-mdr1	43	2e-05
CS2400000 1FF**E94**	*MFS*	XM568393 *Filobasidiella neoformans* multidrug transporter	43	7e-11
		ACH487683 *Acremonium chrysogenum* cefT multidrug resistant protein	40	6e-05

Protein sequences identified by the two methods mentioned above were checked using the BLAST P algorithm back against the EST protein database to ensure that the four identified EST fragments were still showing the similarity to ABC and MFS transporters for which they had been annotated (Table 3.6). For example, to find MFS genes, the EST database was compared for similarity against three known MFS drug transporters (from *A. fumigatus*, *B. fuckeliana* and *G. pulicaris*). In all five EST sequences were identified with similarity to MFS transporters, both E54 and E94 had been selected by searching the annotations of the EST database but 0C9, 168 and 809 were also isolated. The e-values for the MFS genes in particular were very low (Table 3.6), and so the sequences' identity was verified once again with the BLASTX algorithm. This confirmed that among the suspected MFS sequences, only MFS E94 and MFS E54 could be considered to show similarity to transporter genes and likewise from the ABC identified sequences only ABC D5A and ABC FBE were chosen.

The four EST fragments (Table 3.5) were chosen as putative transporter gene fragments that would be analysed and sequenced. Two ABC genes, now called ABC D5A, ABC FBE, and two MFS genes, now called MFS E54 and MFS E94

3.1.6 Summary of ABC and MFS Transporter Gene Identification

3.1.6.1 MFS Transporters: Low Sequence Conservation

All multicellular organisms contain transporters of the ABC and MFS classes. The substrates of these transporters are highly varied and sequence homology is limited to selected regions that carry out specific functions. In the case of MFS transporters, no single amino acid motif is held in common between all members of the superfamily. The sub-families may be recognised by a certain motif but even this motif may be degenerate with only 2 or 3 amino acids recognisable for all genes in this class (Figure 3.2). This makes it especially difficult to design degenerate primers to MFS genes. Conserved regions could not be identified across all the MFS genes included in the alignments. Degenerate primers could not be designed with an optimal degree of degeneracy, nested primers also could not be designed as the regions of conservation were too limited. Consequently, the cloned PCR fragments recovered were not derived from MFS transporter genes.

3.1.6.2 ABC Transporter Identification: Success in Nesting

The degenerate ABC primers were designed to the ATP-binding regions from ABC multidrug transporter genes. These motifs are present in all sub-families as functional structures and contain the Walker A and B motifs and the ATP-binding region. The first set of primers designed to these motifs did amplify and identify ABC sequences, but the e-values were poor and the results also contained matches to other types of genes. Nesting of the primers abundantly increased the number of putative ABC gene fragments cloned.

Table 3.6. BLASTX results of known transporters against the EST database, confirmation of the database annotation

ABC Transporters	*N. lolii* EST	% ID	e-value	BLASTX hits to *N. lolii* ESTs and e-values
CAF32148 ABC transporter, putative *Aspergillus fumigatus*	101004CS2400000 1FFD5A	62	e-109	*Magnaporthe grisea* abc1 transporter (2e-88)
				Neurospora crassa hypothetical protein (6e-89)
	100304CS2100000 1FFFBE	29	2e-10	*Aspergillus nidulans* multidrug resistance protein (1e-55)
				Emericella nidulans ABC-transporter (2e-55)
AAB86640 ABC1 transporter; ABC-type ATPase *Magnaporthe grisea*	101004CS2400000 1FFD5A	76	e-108	*Magnaporthe grisea* abc1 transporter (2e-88)
				Neurospora crassa hypothetical protein (6e-89)
	101004CS2400000 1001B4	47	6e-01	*Ashbya gossypii* (yeast) hypothetical protein (2e-07)
	100304CS2100000 1FFFBE	29	5e-09	*Aspergillus nidulans* multidrug resistance protein (1e-55)
				Emericella nidulans ABC-transporter (2e-55)
CAC40023 ABC-transporter *Gibberella pulicaris*	101004CS2400000 1FFD5A	53	2e-89	*Magnaporthe grisea* abc1 transporter (2.00E-88)
				Neurospora crassa hypothetical protein (6e-89)
	101004CS2400000 1FF97B	52	0.57	*Neurospora crassa* hypothetical protein (1.00E-64)
	100304CS2100000 1FFFBE	28	3e-08	*Aspergillus nidulans* multidrug resistance protein (1e-55)
				Emericella nidulans ABC-transporter (2e-55)
MFS Transporters				
AAW03302 GliA *Aspergillus fumigatus*	101004CS2400000 1000C9	35	1.2	*Neurospora crassa* hypothetical protein (1e-75)
				Ashbya gossypii (yeast) (1e-63)
	100304CS2100000 1FFE94	38	0.7	*Acremonium chrysogenum* multidrug resistant protein (2e-04)
				Neurospora crassa hypothetical protein (1e-35)
AAF64435 major facilitator *Botryotinia fuckeliana*	101004CS2400000 100168	35	1.1	*Neurospora crassa* hypothetical protein (1e-36)
				Neurospora crassa probable translation initiation factor sui1 (1e-33)
CAB69830 mfs-multidrug-resistance *Gibberella pulicaris*	100304CS2100000 1FFE94	38	0.7	*Acremonium chrysogenum* multidrug resistant protein (2e-04)
				Neurospora crassa hypothetical protein (1e-35)
	101004CS2400000 1FFE54	44	4e-07	*Aspergillus fumigatus* putative drug resistance protein (3e-07)
				Neurospora crassa hypothetical protein (1e-07)
	100304CS2100000 1FFE94	38	0.7	*Acremonium chrysogenum* multidrug resistant protein (2e-04)
				Neurospora crassa hypothetical protein (1e-35)
	101004CS2400000 1FF809	33	0.94	*Neurospora crassa* clock-controlled protein 6 (5e-18)
				Saccharomyces cerevisiae cell wall protein (3e-10)

Table 3.6 Legend. BLASTP Results of Known Transporters against the EST Database, Confirmation of the Database Sequence Annotation (page 69)
The results of transporter gene identification from an EST database of *N. lolii* Lp19 sequences. The transporters in the first column (both ABC and MFS) were known transporter genes that were used in the alignments in Section 3.1.3-4. The complete sequence of gene products was compared against the *N. lolii* EST database using the BLASTP algorithm. The EST accession names, in the second column, are the results of this search, and following is a listing of the percentage amino acid identity, length of the aligned sequence and the e-value for this result. The final column refers to the second step in the identification process where the identified *N. lolii* ESTs were re-aligned with external databases using BLASTP algorithms. Top BLASTP results and their e-values are listed here

Eighty-eight fragments were recovered and sequenced from the nested PCR experiment. The cloned fragments were grouped into similar sequences using Align X (Vector NTI). The sequences clustered into 5 groups (Figure 3.13), of which five matched ABC transporters, and one group contained fungal EF-3 (elongation factor 3) sequences and a hydroxy/aromatic amino acid permease (HAAAP). These other proteins also contain ATP-binding sites which have the same amino acid motifs as ABC transporters. Four of the five putative ABC sequences had primers designed to them and further identification of these genes will continue in Reverse-Transcriptase PCR experiments (Section 3.2.4).

Overall if the primers are too degenerate, the products are non-specific. Although decreasing degeneracy in the primers pairs had a big impact on selective amplification of ABC transporter genes, nesting of degenerate primers was the most effective means of amplifying ABC transporters from the genome. The four genomic ABC Transporters identified will now be called gABC 4a (GA3 4a), gABC 4c (GA3 4c), gABC 4e (GA3 4e) and gABC 4g (GA3 4g) to reflect they are genomic sequences identified as putative ABC transporters as opposed to ESTs and to indicate the well number that was allocated to distinguish them.

Four further transporter genes were also identified from an *N.* lolii EST database. Two of these were ABC genes and two were found to be MFS genes. Although degenerative PCR was not a suitable method for cloning MFS genes from the genome, two MFS EST sequences were available for further analysis. Complete Sequence for all of the eight transporter gene fragments, genomic and EST is shown in Section 3.2.2. In this section the fragments will be analysed further and the designing of primer pairs for use in RT-PCR experiments is discussed.

3.2 Analysis and Expression of *N. lolii* ABC and MFS Transporters

3.2.1 Semi-Quantitative Reverse-Transcriptase PCR of putative transporter genes

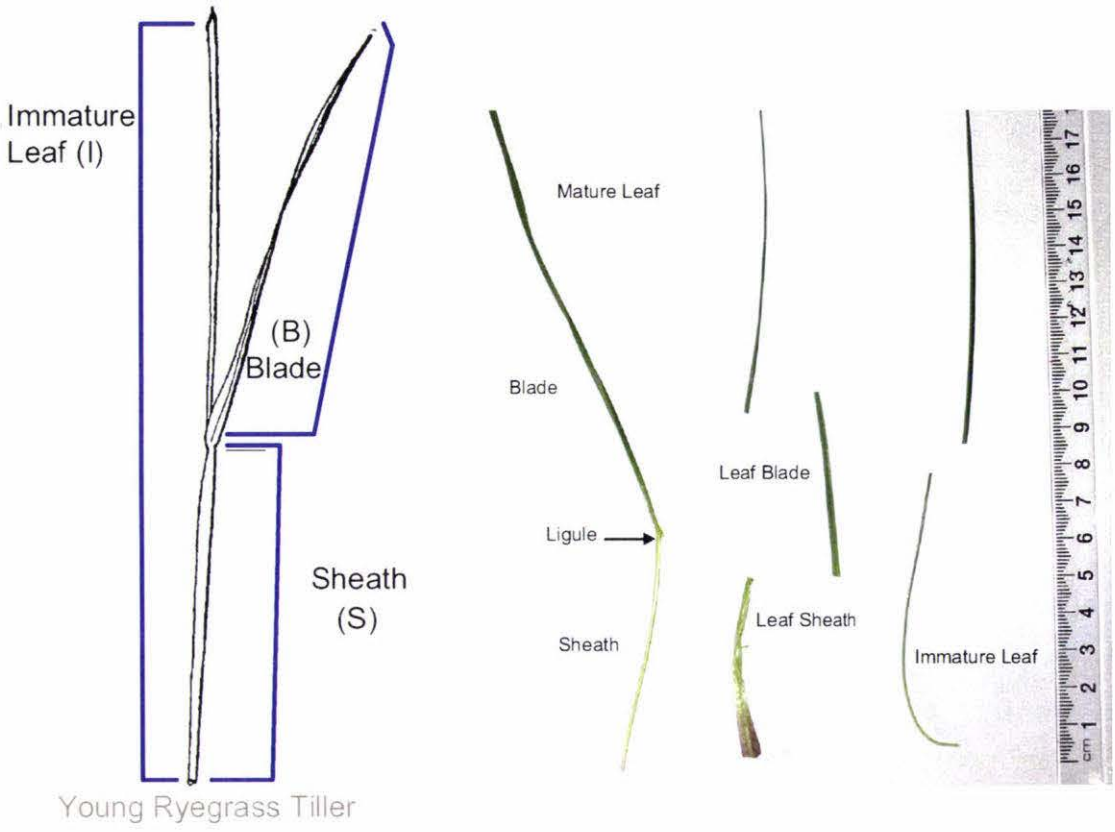
The role of transporters in moving sugars, toxins and signalling molecules could make them vital in symbiotic relationships, such as the interaction between *N. lolii* endophytes and their grass hosts. Expression studies using RT- PCR could suggest a role in symbiosis, if any of the eight transporters shown selected from Section 3.1 are up-regulated *in planta*, and will indicate if the putative genes are expressed at all.

The genes were compared across different tissue types within endophyte-infected plants and the expression was compared both in culture and *in planta*. Looking at expression of the genes within different tissues of the grass host could indicate roles in different endophyte growth stages. Endophyte colonisation is most dense in the leaf sheath, the more mature portion of the leaf (Figure 3.15). The leaf blade is moderately colonised and very little endophyte infection is seen in the immature developing leaves, as the endophyte grows along with developing tissue. The expected result is that expression of the transporter genes would be greater in the leaf sheath due to an increased proportion of fungal biomass in this tissue, unless the genes are involved in the growth and developmental stages of the leaf. The blade should also show high levels of expression, secondary metabolite production is proposed to take place in these matured plant sections (Rowan, 1993).

The experimental plant material was sectioned into sheath tissue, blade tissue and immature leaves (Figure 3.15). The grass tillers from the endophyte-infected or the endophyte-free (control) plants were cut at the base. Figure 3.15 shows the way in which grass tissue was collected, with the immature leaf removed from within the coiled outer leaves and 8 cm measured from the base and collected. The mature leaves were separated into blade and sheath by cutting at the ligules which is a distinctive segmented join between the two tissues. All of the sheath material was collected and the blade material was collected for up to 5 cm from the ligule cut.

The ryegrass plants were mature plants that had been grown in glasshouse conditions. The plants had been manually infected with the *N. lolii* Lp19 strain of endophyte when they were young seedlings. Before this study was initiated they were rechecked to ensure endophyte infection (or lack of it in the case of the controls) by microscopic examination of several tillers. The cultured material was grown on media and kept at 25 °C (see methods Section 2.1.2).

Figure 3.15 Tissue Types for Expression Studies.



The drawing shows one tiller of ryegrass as was used in this experiment. The mature leaf is divided into blade tissue and sheath tissue. The sheath tissue wraps around the immature leaf as it grows from the meristem at the base and protects the developing leaves. The photograph shows the dissection of a tiller into the respective blade, sheath and immature tissues as explained in the text.

3.2.2 Analysis of Transporter Sequence Fragments for RT-PCR Primer Design

Primers were designed to each of the eight sequences for RT-PCR experiments. Care was taken to ensure the primers were specific for the amplification of their unique sequence. The fragments had all been amplified from the same pairs of degenerate primers and so the nucleotide fragments were aligned to avoid designing primers to nucleotide fragments conserved between the fragments.

The amino acid sequences of the gene fragments identified as transporter genes in the *N. lolii* genome or isolated as an EST from the *N. lolii* sequence database, were searched for any recognisable motifs and compared with one another. Since the GA 300 fragments identified in Section 3.1 had been amplified using the same primers; each of them should cover the same region of different genes. This is seen for fragments gABC 4g, 4a and 4c since they all contain the Walker B motif in the early part of their sequence (Figure 3.16a). However sequence GA3 4c does not have sequence identity with the Walker B motif but the Walker A motif instead and so it shows far less similarity with the other three genomic sequences.

The EST sequences identified from the *N. lolii* database were slightly longer than the genomic ABCs identified and do not have any similarity between their nucleotide sequences. The ABC transporter ESTs are close to 800 bp each and the MFS transporters are approximately 500 bp (Figures 3.16 b, c). The EST fragment ABC FBE contained an entire ATP-binding region with the Walker A and B motifs and the ABC motif all recognisable. The remaining EST sequences ABC D5A, MFS E54 and MFS E94 did not have distinguishable transporter motifs. Unique primers were designed to each sequence for RT-PCR and are shown in Figure 3.16.

Figure 3.16 Genomic and EST Transporter Sequences from *N. lolii*

Figure 3.16a. *N. lolii* Genomic ABC Sequences

gABC 4g

Primer 4g_R2 — **Walker B**

```
1    ACGAT CGGTG TGGAG CTGGC CGCCA AACCC CCCTT GCTAC TCTTC GTGGA

51   CGAGC CTACC TCTGG TCTGG ATTCG CAAAC ATCTT GGGCC ATCCT GGACC

101  TTCTA GAGAA GCTTA CCAAG GCTGG CCAGG CTGTT CTCTG CACCA TACAT

151  CAACC CTCCG CCATG CTTTT CCAGC GTTTC GATCG GCTTC TCTTC TTGGC

201  CAAGG GTGGA AAGAC TGTGT ACTTT GGCGA TATTG GCGAG AAATT CCAAG
                                                 Primer 4g_F2
251  ACGAT GACGA GCTAT TTTGA GCGAA AATAG TGGAC TCACT TGCCC TGAAA
```

gABC 4a

Primer 4a_R2 — **Walker B**

```
1    ACCAT TGGCG TGGAA CTAGC CTCCA AGCCC GAGCT CCTCA TGTTC CTCGA

51   CGAGC CTACA TCCGG CCTCG ACTCG GGAGC AGCCT TCAAC ATTGT CCGCT

101  TTCTG CGAAA GCTCG CGGAC GCGGG ACAAG CAGTC TTGTG CACCA TCCAC

151  CAGCC ATCCG CCGTT CTATT CGAGC ACTTT GACGA ACTAC TACTG CTCAA

201  GTCCG GAGGA CGGGT AGTCT ACCAC GGCCC GCTGG GAGAA GACAG CAGAA
                                              Primer 4a_F2
251  CCCTC ATCGA CTACC TCGAG TCCAA TGGCG CGTTT CAAGT GCCCC CCAGA
```

gABC 4c

Primer 4c_R2 — **Walker B**

```
1    ACAAT TGGGG TGGAG CTAGC GGCCA AACCG AAGCT TCTAC TATTT CTTGA

51   CGAAC CAACT TCTGG TCTGG ACTCT CAGAG CTCCT GGGCG ATCTG TGCCT

101  TCCTT CGAAA GCTGG CCGAT TCTGG CCAGG CTATC CTGTG TACCA TTCAT

151  CAGCC CAGCG CTATC TTGTT CCAAG CCTTT GATCG CCTGC TATTC CTTGC

201  CAAAG GTGGA AAAAC GGTTT ACTTT GGCAA CATTG GCGAC AACTC GCGCA
                                              Primer 4c_F2
251  CCATG TTGGA CTACT TTGAG CAGAA GCGGG GCACG AAAAT GCGAT GACAA

301  GGAAA ACCCA GAGGA GTGGA TG
```

gABC 4e

Primer GA3a_4eR — **Walker A**

```
1    GCGCC AATGT CGCCT TTGTT GGCAC CACAG GTTCT GGCAA GAGCA CCATC

51   GTCTC GCTGA TTGAG CGGTT TTATG ACCCA ACCAG TGGGA GGATA CTTGT

101  GGACT CCAAA CCTGT TACGT CGCTC CGCTT GTCCG AATAT CGAAA TGCAT

151  TGGTT TGGTC AGCCA AGAGC CCACC CTGTT CAACG GCACC GTCAA GATGA

201  ATTTG ACGAT TGGTT TGGAC GAAGA GGGCC AAACA CCACC CACTA ATGAG
       Primer GA3a_4eF
251  GAAAT CGAGA ATGCG TGTCG
```

Figure 3.16b. *N. lolii* **ABC EST Sequences**

ABC FBE

```
1      GGCAA TATCA CGTTC GACTC GGTTA GCTAT TCATA CCCTT CCNNN NNNNN

51     NNACC TGGCG TTGGA CAATG TCAGC TTCAA CATTC NNNNN NGCGC CAATG
       Primer ABC FBE_F                    Walker A
101    TCGCC TTTGT TGGCA CCACA GGTTC TGGCA AGAGC ACCAT CGTCT CGCTG

151    ATTGA GCGGT TTTAT GACCC AACCA GTGGG AGGAT ACTTG TGGAC TCCAA

201    ACCTG TTACG TCGCT CCGCT TGTCC GAATA TCGAA AATGC ATTGG TTTGG
         Primer ABC FBE_R
251    TCAGC CAAGA GCCCA CCCTG TTCAA CGGCA CCGTC AAGAT GAATT TGACG

301    ATTGG TTTGG ACGAA GAGGG CCAAA CACCA CCCAC TAATG AGGAA ATCGA

351    GAATG CGTGT CGCTC AGCCA ACATT CATGA ATTCA TTACC AGCCT TCCGG

401    ATGGC TACGA CACGG AAGTT GGCAG TAGAG GCAGT CAGCT CAGTG TCGGA
           ABC
451    CAAAA GCAGC GTGTG GTGCT CGCCC GGGCC CTGCT TCGGC GACCT AAGGT
           Walker B
501    TCTGT TGCTC GACGA GGCGA CGTCT GCGCT GGATT CGCAA TCAGA AGCCA

551    GCATT CAACA GGCGC TTGAG CTGGC TAAAG AGGGT AGGAC AACAA TCACG

601    ATTGC GCATC GACTG AGTAC AATTG TAAAA GCGGA CAAAA TTTAC GTCAT

651    GGGTG ACGGC AAAAT TGTCG AAGCG GGAAC ACACG CTCAG TTGAT GGCCA

701    AGAAG GGTGC TTATC ATGGA CTGTA CGTGG CAAAC AAAAN NGGGC AAACT

751    CTGTA AATGT TGTGA ATGCT GTANN TGTGG CAGAT TATTA TTGTA CCAAT

801    TC
```

ABC D5A

```
         Primer ABC D5A_F
1      CGTTG TATGA GGCTC GTGAG CGCCC ATCCA AAGTA TACTC GTGGA NNGTG

51     TTCAT GCTCT CCCAG ATCAT CGTCG AACTG CCCTG GAATG CCTTG ATGGC

101    CGTCA TCATG TACTT CTGTT GGTAC TATCC TGTCG GTCTC TACCG CAATG

151    CCATC CCCTC CGATG CCGTC ACCGA ACGTG GATTC CTCAT GTTCT TATTT

201    CTGCT CATGT TCATG CTCTT CACGG GTACC TTCTC GACCT TCATC GTCGC
         Primer ABC D5A_R
251    TGGGT TCGAG ACTTC CGAAG CAGGT GGTAA CCTTG CCAAC TTGAT GTTTA

301    CCCTC TGCCT GATCT TCTGC GGTGT TCTCG CCTCA CCTGA CAGCC TCCCC

351    CGTTT TTGGA TTTTC ATGTA CCGCG TCTCT CCGTT CACAT ACATC GTTTC

401    AGGCA TGTTG TCTAC CGGTG TGGCC AACAC CGAAG TCGTC TGCGC TTCAA

451    ACGAG CTGCT AAAGT TCCAG CCGCC TGCCA ACCAA ACCTG CATCG AATAC

501    ATGCG AGACT ACATT TCAGT GTTTG ATGGC AAGCT TTCAG ACGAC ACCGA

551    GAGAG GCGCT TGCCA TTATT GTCCC ATCAG TGACA CGAAC AAGTT CCTGG

601    CTAAC GTGAG CAGCG ACTAC AGCGA GCGCT GGCGC AACTT TGGTT TGCTT

651    TGGGT GTACG TCATT TTCAA CATTG CTGCC GCCAT GTTTG TATAC TGGCT

701    TGTTC GAATG CCCAA GAACA AGCTT GGCGG AAAGA AGGAA AAGAA GGAGT

751    GAATT GGCAA GTT
```

Figure 3.16c. ***N. lolii*** **MFS EST Sequences**

MFS E94

```
1     TTCGGAATAC GCCGATGGCT CTATTCAGTG GCTTTGTACT TGCTGGGACT

51    GGCGCTGGCC CATTGGTTGG CTCAGCCTTC ATGGAAACAC TTGGGCAAAC
                                          Primer MFS E94_R
101   AACGTTCGCT TGGAAATGGA CCTTCTGGCA CCAGGTCATA ATGGACGGCA

151   TGCTTTTTAT TTCCTTTATT TTTTTGTTCA AGGAGAGTCG GGCCTCTGTC

201   TTGCTGAGCA AAAAGGCAAA AAAGCTGAAT CAGTGGTACG AAGAGTTGGA

251   AGCGAACGGT GTGTTTGGAA TGTACTTGCT AGGCTCTTCG ATGCCGATTG

301   CTGCACGATC GGCATCCTCC AGCGATACCT GCTTGTCCTC TGGTGCAAAT
                                        Primer MFS E94_F
351   GACCTGGAAA AAGACTCTGT TACTGCATCC GTGAAGCGAG AAAGTCTTCT

401   GCGAATTCGA TGGGTCGTCG AGGAAGATGA AACCCGACCT TCTGTGCTCC

451   AAATGATGGC CACCTCTGTA AGACGCCCTT TTCACATGCT TTTCACCGAA

501   CCAGNNNCCT TTTNNNNNNC CCTCTGGGCA GCTTTTTCAT GGGCAGTACT

551   TTATCTAT
```

MFS E54

```
       Primer MFS E54_F
1     TCGGATTCAC TGGGCAACTT CTATACTGGC TTTCCTCGCG CTAGTTTGCA

51    TGCCTTTTAC GTTCATCATG TATCGATATG GTGGTCCGTT GAGGATGAAA

101   TGCGAATATG CTTTTGAAGC GGCCGAGATG ATGAAAAAGA TGCAGACACA

151   ACAAACAACG ACCGCGTTTA TCGCAGACTC GACTCAGTGC ATTTCTGTCC

201   GTGCAGTTAT ATGCTCGCAG AAACAGTGTG TGCCGCTGTG AAAGACGCCG
          Primer MFS E54_R
251   CGCGGCTTGC AGAGCATCTA CTACTAATTC CCAAGCCTGT CGGTGGCTAC

301   AGGCGACGAA GTTGTTTCAA GGCATCGCAA TGTACGTCTT CTGGCCGGGG

351   CTACAGTAAA TCTCGGCATT CGGTTCGTCA AACGTTCGTA TAAAAGGGTT

401   TAGCCGAGAT TTGAGGAACA ATACTCCGTA CCTGGTAGCT TAGTAGTACG

451   GATATACCCG CGCTTGCTCC GTACAAGTGG CGGAATGTAT AAAGAATACA

501   TGAGAGCAAG ATTCATGCTG CTTTGTAA
```

Figure 3.16 Sequences of the four chosen putative genomic ABC gene fragments and the four EST genes, as well as the positions of the primers designed to them. Any transporter motifs seen in the sequences are indicated in bold.

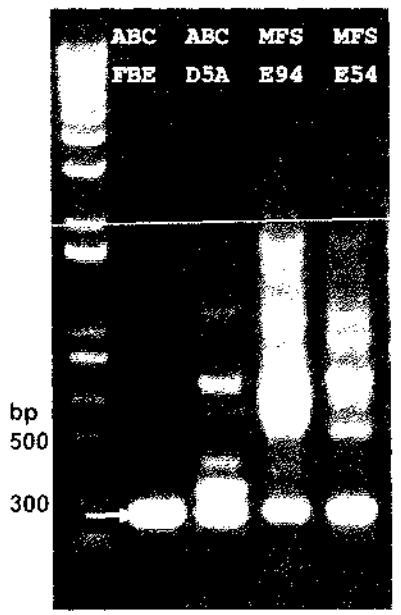
3.2.3 Design and PCR Optimisation of ABC and MFS Primers for RT-PCR

Specific primers had to be designed that would amplify only the target gene fragments and resolve clearly on a gel for reverse transcriptase experiments. The reverse transcriptase primers were all chosen to amplify products of 300bp.

3.2.3.1 Design and Optimisation of EST Primers

Primers were designed to the four EST sequences (Section 3.1.6) to be used in RT-PCR and for identification of genes in BAC and Lambda libraries. The EST fragments were longer than the genomic isolated fragments, however high quality sequence for most of the EST sequences did not extend much beyond 300 bp (Figure 3.16b, c). The primers were designed to amplify the whole fragment. To determine whether all the primer pairs were working, a test PCR of 2mM Mg^{2+} with a 48.5°C annealing temperature was performed. The PCR results in Figure 3.17 show all primer pairs tested amplified a product but most required optimisation to prevent amplification of non-specific DNA. The ABC FBE primer pairs showed a specific band at just under 300 bp, so no further optimisation was required for this pair. PCR optimisations for RT-PCR primers were performed as for the degenerate primers (Section 3.1.3).

Figure 3.17 Pre-optimisation for EST Primer Pairs.

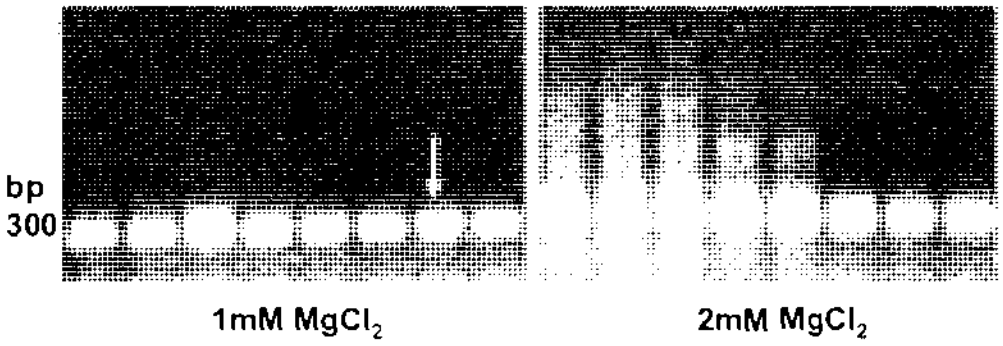


Lane 1 shows primer pair ABC FBE has no need of further optimisation, the product is 270 bp. Lanes 2, 3 and 4 show PCR with the ESTs D5A, E94 and E54 all of which show a product at the expected size but also contain non-specific products.

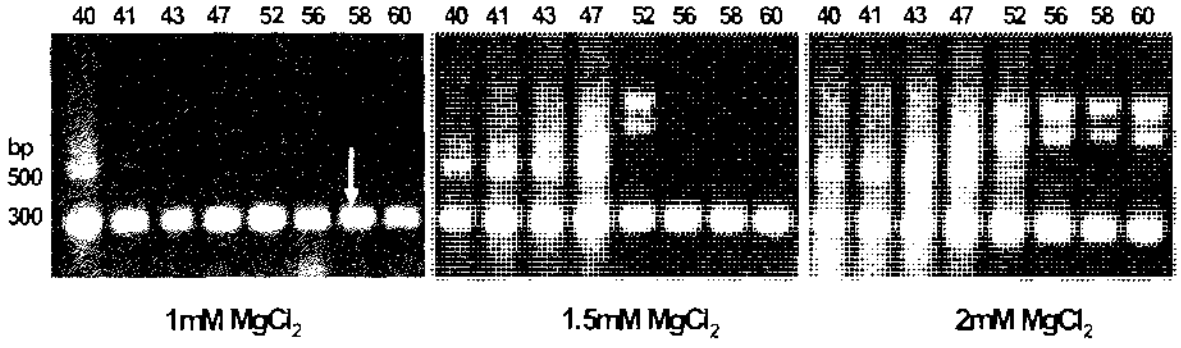
The expected band sizes for each primer pair were: ABC D5A (269 bp), MFS E54 (272 bp), MFS E94 (273 bp) (Figure 3.18). Optimal conditions were chosen for each primer pair by identifying which lanes contained the cleanest and strongest band for each. An annealing temperature of 58.6°C with 1mM $MgCl_2$ was subsequently selected for all primer pairs.

Figure 3.18 EST Primer Pairs PCR Optimisations.

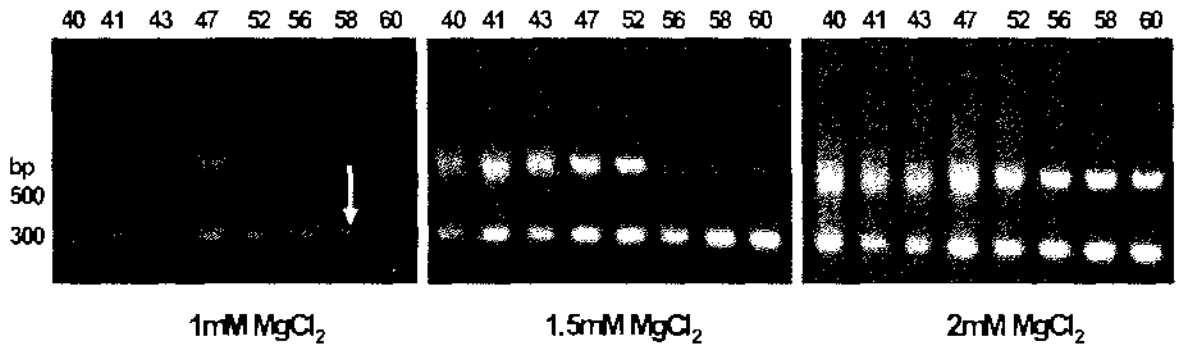
3.18a. ABC D5A



3.18b. MFS E54



3.18c. MFS E94



The optimisations of PCR primers were run with eight different annealing temperatures between 40-60°C as shown along the top of the gel photographs (3.18a, b and c), and at the $MgCl_2$ concentrations, 1mM, 1.5mM and 2mM. The arrows indicate the optimal PCR conditions to obtain a single product.

3.2.3.2 Genomic Primers Design and Optimisation

Primers for the genomic ABC fragments were designed to amplify as large a product as possible, so the expected product sizes are also approximately 300 bp. Each primer set had to be unique so as not to amplify several transporters in one PCR reaction. The four genomic ABCs were aligned in vector NTI Clustal W for this reason. The alignment shown in Figure 3.19 identifies differences between the sequences, regions where the order of nucleotides are distinctive from the other sequences in the alignment and hopefully other sequences within the genome. The primers are highlighted in blue (forward and reverse) in Figure 3.16a, which shows the complete sequence for genes and the relation of the primers to ABC motifs identified in each fragment.

Figure 3.19 Nucleotide Alignment of Genomic ABCs

Clustal W alignment created in Vector NTI comparing the four identified genomic ABCs. This alignment shows the differences between the fragments.

To amplify ABC transporter gene fragments isolated by degenerate PCR, primers were designed and optimised as described in Section 3.2.3.1. The $MgCl_2$ and annealing temperature conditions were kept as similar as possible for each of the primer pairs. For some of the fragments two sets of primers were ordered and trialled (results not shown): the pair giving a single band of the expected size was then chosen for the later experiments. The optimised primer pairs, four designed from genomic sequence and four designed to ESTs, have their optimal conditions for amplification stated in Table 3.7 below.

Table 3.7 PCR conditions for primers used in expression studies.

Primer Pair	$MgCl_2$ conc. (mM)	Annealing Temp. (°C)
gABC_4a	1	58.6
gABC_4c	2	58.6
gABC_4e	1	58.6
gABC_4g	2	56.0
ABC FBE	2	56.0
ABC D5A	1	58.6
MFS E54	1	58.6
MFS E94	1	58.6

3.2.4 Reverse-Transcriptase Expression Analysis

3.2.4.1 RNA Extraction

As described in Section 3.2.1 (Figure 3.15) each ryegrass tiller was dissected into three different tissue types: blade tissue, sheath tissue and immature leaf tissue (the young leaf as yet undifferentiated into blade and sheath). Tissue extracted from plants that were infected with endophyte Lp19, or from endophyte-free ryegrass, was harvested and immediately placed into liquid nitrogen. The material was stored at -80°C until required. Total RNA was extracted from these tissues (Section 2.4) and from *N. lolii* Lp19 cultures. The RNA extraction generally gave good yields of the tissues, the highest yield came from the endophyte grown in cultures and the sheath tissue gave a 10-fold lower yield than the rest of the plant tissues.

3.2.4.2 cDNA synthesis

cDNA for all tissue types was synthesised using the ThermoScript™ (Invitrogen) RT-PCR system (Section 2.4.2.). The sheath material, which had a low RNA yield, had a lesser amount of cDNA synthesised as the concentration of the RNA was too low to allow enough to be added to the reaction. However the resulting product still worked well in the PCRs. To control for contamination of RNA with genomic DNA a 'no reverse transcriptase control' was prepared alongside the reverse transcriptase reactions. This control contained the same amount of RNA starting material, but no cDNA was synthesised in these reactions. The no RT controls were run alongside their cDNA counterparts as a negative control for genomic DNA contamination.

The endophyte-free material was a control to ensure that any expression products amplified in the RT-PCR were products derived from the endophyte and not the host plant. The endophyte-free tissue samples were checked alongside the endophyte-infected ryegrass tissues in preliminary experiments using actin as a housekeeping gene to confirm that the uninfected plants were endophyte-free.

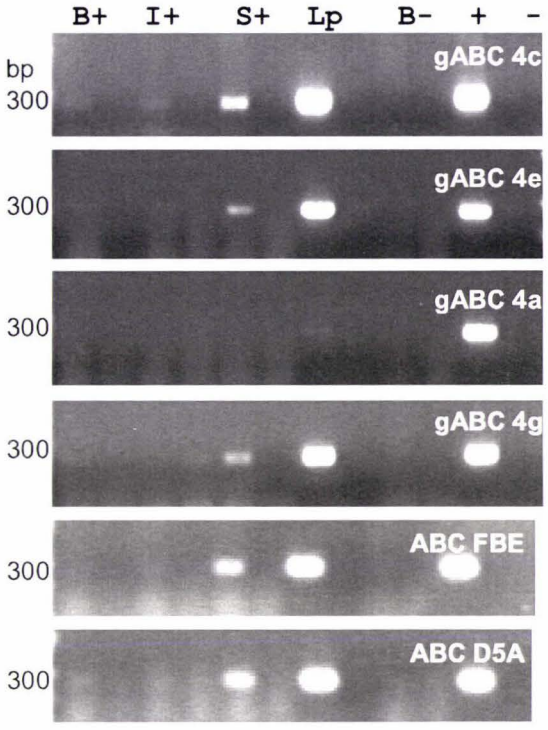
For comparing expression of transporter genes in endophyte growing in culture versus *in planta*, the relative biomass of endophyte *in planta* is taken as approximately 1% (Panaccione *et al.*, 2001, Young, 2006). The cDNA was diluted to levels where the expression of an endophyte house keeping gene, in this case actin, was seen to be similar both *in planta* and in culture. Several dilution series were attempted to ensure that these levels were adjusted correctly (data not shown). For the experiments described below, the Lp19 cDNA derived from *in vivo* cultures, was diluted 1:100 for comparison with the *in planta* tissues.

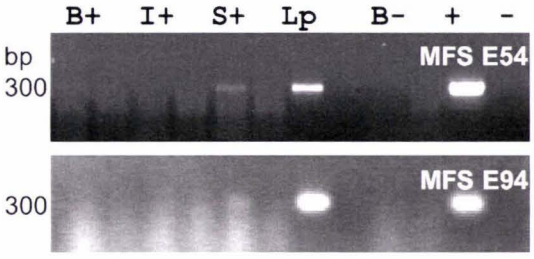
3.2.4.3 cDNA Expression of 8 Different Transporter Genes

RT-PCR reactions were set up with all eight primer pairs against the cDNA tissue samples to compare expression of transporters in different parts of the plant. This is a semi-quantitative plus/minus experiment as the primer pairs are not normalised to enable comparison of one gene to another. The RT-PCR experiments used optimal primer conditions for $MgCl_2$ and annealing temperatures as shown in Table 3.7. All cDNA samples in the results are compared to their no RT controls as a negative control

for contaminating DNA and the experiment was repeated to ensure the results were consistent.

Figure 3.20 Reverse Transcriptase Tissue Expressions





Primer pair	B+	I+	S+	Lp19
gABC_4c	+	+	++	+++
gABC_4e	+	+	++	+++
gABC_4a	-	-	-	+
gABC_4g	-	-	+	++
ABC FBE	-	+	++	+++
ABC D5A	+	-	++	+++
MFS E54	+	-	+	++
MFS E94	-	-	+	++

The eight sets of primers were run against the same tissue samples; infected blade (B+), infected immature leaf (I+), infected sheath (S+), cultured Lp19 1:100 dilution (Lp), endophyte-free blade tissue (B-), a positive control Lp19 genomic DNA (+) and a negative template free control (-). All the cDNA reactions were run alongside their no RT controls. The presence/ absence table indicates expression as a + and no expression as a -, increasing levels of expression are shown by more plus symbols.

All of the genes are expressed in cultured endophyte, although gABC 4a is barely visible (Figure 3.20). With the exception of gABC 4a expression levels for all of the transporter genes appear very similar and none of them appear to be expressed exclusively or to be highly up-regulated *in planta*. The strongest levels of expression for all of the genes *in planta* are seen in the sheath tissue, which is the host tissue most heavily colonised by endophyte presence. Expression was seen in the leaf blade for gABC 4c, 4e, ABC D5A and MFS E54 using a UV transilluminator, although it is faint and hard to see on the gel photos.

Of the ABC genes analysed gABC 4c and gABC 4e also show low expression in both the blade and the immature leaf tissues. gABC 4g is not visibly expressed in leaf blade or immature leaf tissues. The EST ABC gene ABC FBE is expressed in immature tissue

but not visible in the leaf blade, while ABC D5A is seen in the leaf blade but not so clearly in the leaf sheath. The MFS ESTs are also strongly expressed in culture compared to the *in planta* levels. MFS E54 is seen in sheath tissue and in blade tissue, but MFS E94 is seen only in the sheath tissue (Figure 3.20).

3.3 Identifying Transporter Genes in the BAC Library

A *N. lolii* Lp19 BAC library was screened to identify BAC clones containing transporter genes to investigate whether two or more of the transporter fragments may be linked to one another in the genome, and possibly be parts of the same gene. This was possible as the EST identified fragments were not from the same regions of the gene sequence as the genomic transporter fragments were. Furthermore, with the gABC primers designed to an ATP-binding region it was possible that two of these motifs could have been amplified from the same gene (N- or C-terminal ATP binding regions). It was also possible that the EST sequences would be from the same genes as each other. The BAC library was also screened with the intention of identifying BACs to use for direct sequencing of the genes of interest.

The *N. lolii* Lp19 BAC library was created for the Advanced Fungal Technologies team, (AgResearch Grasslands, Palmerston North New Zealand) by Amplicon Express (Pullman, WA, USA) (Section 2.5.1). To create the library, genomic *N. lolii* Lp19 DNA was partially digested with *MboI* and cloned into pECBAC1 (Table 2.1). Glycerol stocks of the BACs were stored in four 384 well plates which were then pooled into a single 96 well glycerol plate, each well containing sixteen BACs for easy screening by PCR.

3.3.1 Screening the Pooled BAC Plate

The BAC library was initially screened from the pooled 96 well plate (Section 2.5.1) by PCR, using conditions determined by the primer optimisations (Table 3.7). A positive result was a PCR product at approximately 300 bp (the approximate product size for all the primer sets).

As seen in Table 3.8 many of the BAC pools contain a putative transporter gene with homology to the primers used in this screen. All of the transporters identified so far were well represented in the library with the exception of MFS E94 which was not found. From the pooled 96 well plate of BAC clones, 53 wells gave a 300 bp band to one of the primer pairs, which gives an indication of how frequently transporter genes are represented within the *N. lolii* BAC library.

Primers ABC D5A and gABC 4g amplified DNA from the exact same BAC colonies as each other, they were both found in the same 11 wells. Primer sets gABC 4e and ABC FBE also amplify bands in three of the same wells, however ABC FBE also amplifies products in three other wells. The rest of the primers appear to be designed against independent transporters as they have no wells in common.

Table 3.8 Pooled BAC Screening Results

Primer pair	gABC 4e	ABC FBE	ABC D5A	gABC 4g	MFS E54	gABC 4c	gABC 4a	MFS E94
Well ID	8G	8G	2C	2C	3B	1D	1H	No BAC
Pooled	9G	9G	3B	3B	4E	2A	2F	
BAC plate	10F	10F	3F	3F	5B	3A	3C	
		5B	5D	5D	6C	4F	3D	
		5F	6F	6F	8E	5A	5A	
		8E	7C	7C	9C	5D	5G	
			8A	8A	10A	6A	8E	
			8H	8H	12G	6D	9C	
			10D	10D		7C	10G	
			10E	10E		7G	11E	
			12E	12E		10B	11D	
						10C	12F	
						10D	12C	
						11H		
						12F		

Wells that amplified a positive product (300 bp) for each of the 8 primer pairs from the pooled 96 well BAC Plate. Each positive well represents a mixture of 16 individual BACs which then have to be screened. The samples indicated by the blue font are those that were picked to follow through with.

For each transporter gene one well from the 96 well of pooled BAC clones was chosen for continued screening to identify the individual BAC clone. Wells that were not adjacent to any other positives were chosen to reduce false positives. For the primers identifying possible linked sequences, one representative was picked. To identify the individual BAC clone for each primer pair, a PCR screen was performed using the 16 BAC colonies corresponding to each pooled well. This experiment and its results are

summarised in Table 3.9. The original 384 well BAC plate reference is stated along with the well numbers that correspond to the pooled well chosen from the 96 well plate. The final column indicates which well contained the individual BAC colony with the putative transporter. Pooled wells 8G and well 5D which had been amplified by more than one primer pair were checked against each primer set to ensure that both gave the same positive result.

Table 3.9 Unpooled BAC Screening Results

Pooled Well	BAC Plate	Screened Wells	PCR positive BAC
8G	NL003	N01-N08 M17-M24	M23
5D	NL002	H01-H08 G17-G24	G22
4E	NL002	I01-I16	I14
4F	NL002	K01-K16	K1
5A	NL002	B01-B08 A17-A24	A22

The pooled well column indicates which well was chosen from the BAC 96 pooled screen. The individual BAC that may contain the putative transporter gene is shown on the far left.

Each of the five BACs identified were streaked for single colonies from the glycerol stocks, grown up as overnight stocks and re-checked by PCR. The BAC plasmids were isolated and sequencing was attempted (Section 2.5.1). However although BAC sequencing had been optimised to work for other genes, the transporter genes could not be sequenced in this way. Instead, Lambda libraries were screened and phagemids isolated for sequencing from these (Section 2.5.2, 3).

The BAC library screening showed some gene fragments are linked to a relatively similar place within the genome suggesting that they may be different parts the same gene or different genes located in close proximity. ABC D5A was derived from an EST and did not contain any of the ABC motifs against which the degenerate primers were

designed (Figure 3.12b), however ABC D5A and gABC 4g co-localised to BAC G22 (Table 3.9). Aligning these two sequences using Vector NTI contig express did not create a contig, but an ABC transporter is a large gene (4.5 kb) and the likelihood of the two fragments overlapping if they are parts of the same gene are low. Gene fragments gABC 4e and ABC FBE were also linked to a shared BAC clone, M23 (Table 3.9), although ABC FBE was also amplified in three pooled wells that gABC 4e was not (Table 3.8). This could be a reflection of the size difference between the fragment; the fragment of ABC FBE shows a complete ATP-binding region with a Walker A, Walker B and an ABC motif while the sequence of gABC 4e contains only a Walker A motif. This may indicate a truncated gene sequence for the BAC clones identified with ABC D5A only, that do not contain the Walker A motif that the gABC 4e primers were designed around. Aligning the sequences using Contig Express (Vector NTI) gives an exact match for all 270 bp of the gABC 4e fragment.

3.3.2 Selecting Putative Transporters of Interest

Examination of the transporter genes from the EST database the genomic ABC transporters, for their locations in the BAC library indicated that some fragments likely belong to the same gene or two closely linked genes as they fall on the same BAC clone. For example gABC 4e and ABC FBE are a part of the gene gABC 4e, as determined by sequence alignment. Fragments gABC 4g and ABC D5A belong to the same gene gABC 4g or are linked. The remaining genomic ABCs, gABC 4a and gABC 4c, and MFS E54 appear to be transporters independent of one another. Library screening eliminated the MFS gene MFS E94 from further examination as it was not found in the BAC library or in either of the lambda libraries, leaving only one MFS transporter gene MFS E54 for sequencing.

The reverse transcriptase expression studies were conducted to identify putative transporters that may be important in symbiosis. The genes that were of interest were those that appeared to be up-regulated *in planta*. The expression levels of the putative genes on the same BAC clones were viewed with the idea of them being the same genes. ABC transporters ABC D5A and gABC 4g are linked on the same BAC and had a similar pattern of expression in RT-PCR experiments. Also ABC FBE and gABC 4e showed similar expression patterns in the RT-PCR experiments. The gene gABC 4a was

not highly expressed in any tissue and the gene MFS E94 was not amplified in the BAC and lambda libraries.

Transporter genes of interest were selected for further analysis. The genes were selected on the basis of their BLASTX hits to other known transporter genes, their expression levels *in planta* and also on the basis of BAC clone localisation. This left four transporter genes for sequencing and further studies: MFS E54, gABC 4e, ABC D5A and gABC 4c.

3.4 Distribution of Transporters in Other Endophyte Strains

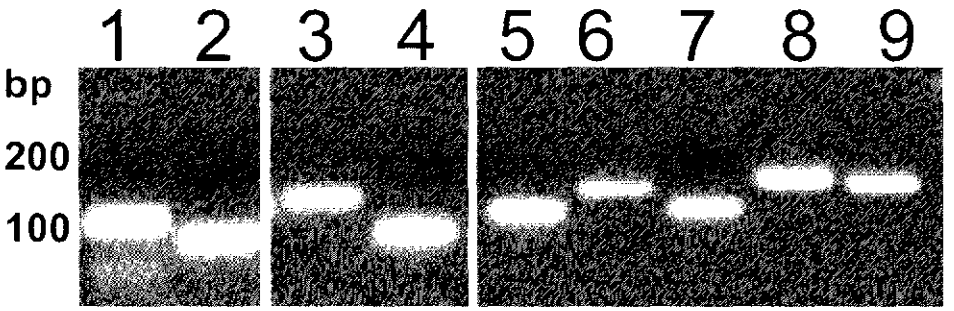
Endophytes live in a variety of host plants that survive in very diverse climates and conditions. Individual endophyte strains have the ability to produce different secondary metabolites from one another; many infect their host species only and may not be able to infect other grass species. ABC and MFS transporter proteins are responsible for the transport of a wide range of substrates; possibly some of these molecules have a part to play in endophyte/plant symbiosis. Screening a transporter gene across a selection of different endophyte strains may indicate a linkage to a class of metabolites if the expression of a particular transporter correlates with the production of the same metabolite in several strains.

The transporters used for this study were identified as genes that gave strong PCR products, were recognised *in planta*, some showed linkage between EST genes and genomic fragments and include both ABC and MFS transporter classes. gABC 4c is a gene of interest because of its high *in planta* expression levels and the fact that it shares sequence similarity and is linked on the same BAC as the EST identified gene ABC FBE. The genomic ABC gABC 4g is similarly chosen for its link on a BAC chromosome to the EST gene ABC D5A. Both MFS genes were considered even though MFS E94 was not found in the BAC library and gABC 4c is continued with because of its consistently high expression *in planta* and in culture, making a total of five for further analysis.

3.4.1 Primer Optimisation

The primer sets used for this distribution study were originally designed for real time PCR. The sets were used for this strain distribution study as they were unique to their own transporter fragment. The primers were designed to amplify an average of 100 bp with 2 pairs designed for most of the sets (Table 2.1). All primer sets were checked at the 60°C annealing temperature for which they had been designed and the results are shown in Figure 3.21. The strain distribution study uses genomic DNA; the positive control for this experiment is the *N. lolii* actin gene (Primers Table 2.1). All other PCR conditions and reagents are described in Section 2.3.1.

Figure 3.21 Evaluation of Small Primer Sets



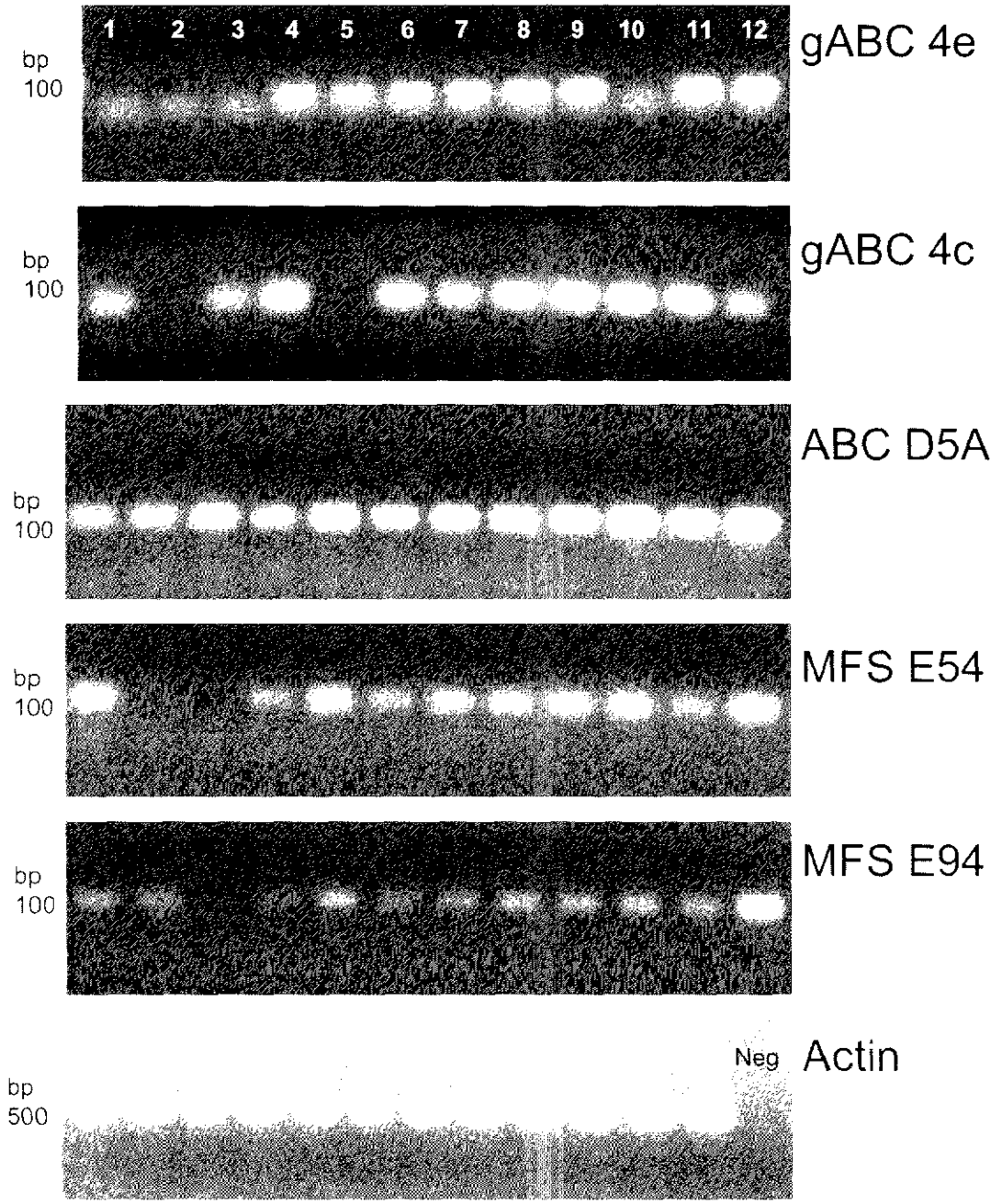
All primers were used with an annealing temperature of 60°C. Primers used in each reaction are as follows; **lane 1** gABC 4c F1/R1, **lane 2** gABC 4c F2/R2, **Lane 3** gABC 4e F1/R1 **lane 4** gABC 4e F2/R2, **lane 5** ABCd5a F1/R1, **lane 6** ABCd5a F2/R2, **lane 7** MFSe54 F1/R1, **lane 8** MFSe54 F2/R2, **lane 9** MFSe94 F1/R1.

3.4.2 Presence of transporters across 12 Endophyte strains

Neotyphodium endophytes are asexual anamorphs of the *Epichloë* endophyte symbionts of temperate grass species. Both *Neotyphodium* and *Epichloë* strains have different ranges of infectivity and specific host plants. Figure 3.23 is a phylogenetic tree for 48 strains, clustered into species groups. The *Neotyphodium* endophytes: *N. lolii, N. coenophialum and N. uncinatum* are found in the grass hosts perennial ryegrass, tall fescue and meadow fescue respectively. While the sexual *Epichloë* species *Epichloë typhina* and *Epichloë festucae* have a broad host range found in *Festuca, Lolium*,and *Koeleria spp* (Schardl, 2001).

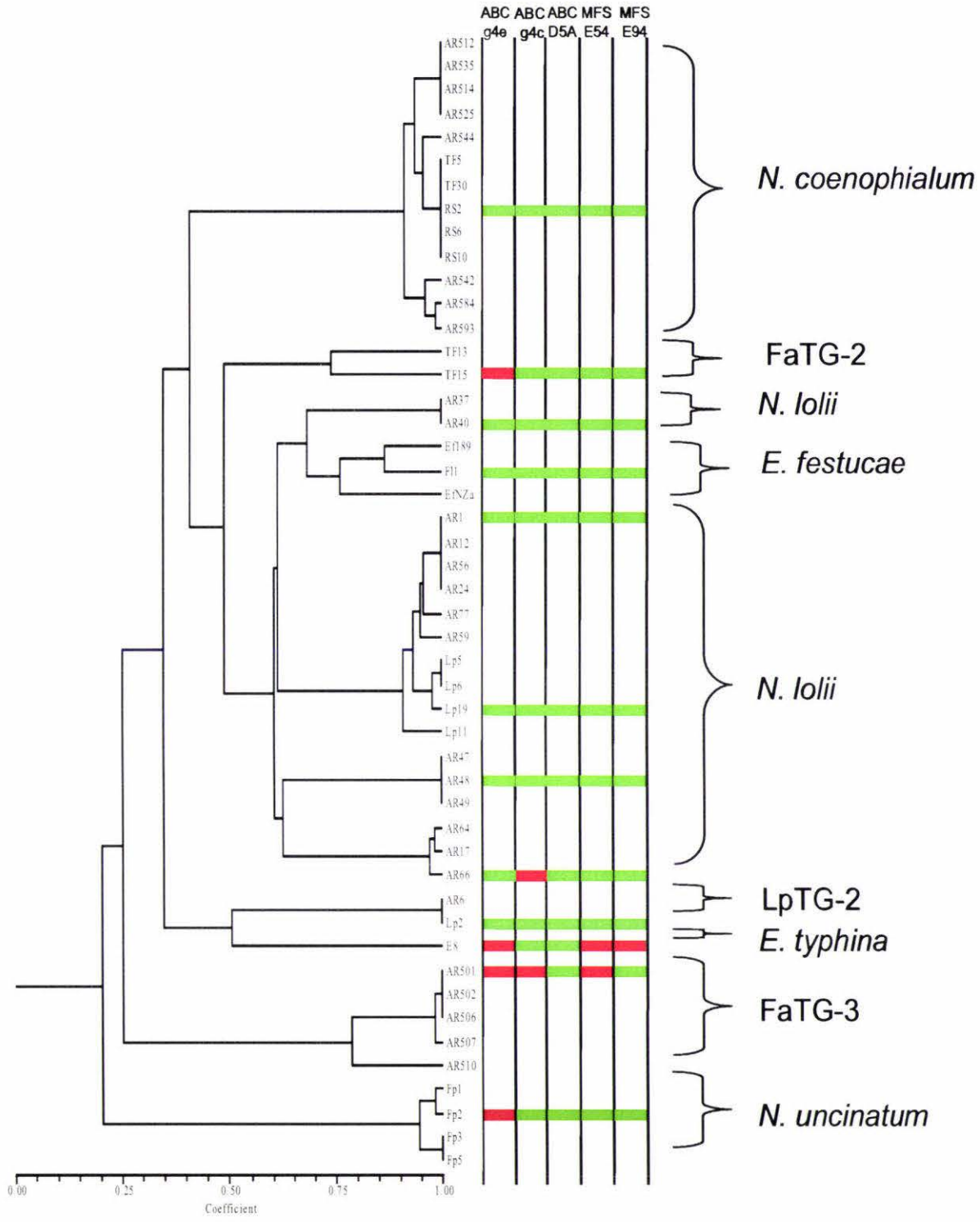
Twelve strains (including Lp19) were chosen for testing with each of the transporter genes as an indication of the presence or absence of the genes within the genomes (Figure 3.22). Shown in Figure 3.23 are the results of AFLP analysis across a range of endophytes, indicating the degree of relatedness between isolates, strains and species. One sample was picked from each cluster to infer the presence of a transporter gene within each group. The genomic DNA of the strains and the phylogenetic tree showing the genome similarity using AFLP markers are courtesy of Andrew Griffiths (AgResearch, NZ.).

Figure 3.22 Presence or Absence of Transporters in Endophyte Strains



Lanes numbered 1-12, represent the strains Fp2, AR501, E8, Lp2, AR66, AR48, AR1, Fl1, AR40, Tf15, Rs2 and Lp19 respectively (details of strains are shown in Table 3.10). The positive control actin also shows a no template control (Neg) in the thirteenth lane. Primer sets used were gABC 4c F2/R2, gABC 4e F2/R2, ABCd5a F2/R2, MFSe54 F1/R1 and MFSe94 F1/R1.

Figure 3.23 Phylogeny of Endophyte Strains and Presence/Absence of Transporters



Phylogenetic Tree Courtesy of Andrew Griffiths

Endophyte species are indicated on the right and the individual strain names at the ends of the tree branches on the left. Presence (green) or absence (red) of the transporter genes is indicated alongside each of the strains tested.

Five transporter genes, three ABC (gABC 4c, gABC 4e, ABC D5A) and two MFS (MFS E54, MFS E94), were screened across twelve different endophyte genotypes (Figure 3.23). Seven of the strains (RS2, AR40, Fl1, AR1, Lp19, Ar48 and Lp2) were shown to hold all five of the transporter genes within their genome. ABC D5A was the only gene present in all twelve of the endophyte genomes. ABC 4e is seen in 8 of the twelve strains making it the most varied in its presence of the five genes looked at. This is the gene that was chosen for complete sequencing and gene analysis.

Strain Lp19 of the *N. lolii* endophyte from which these putative transporters were identified produces three alkaloid toxins, peramine, ergovaline and lolitrem. This strain distribution study notes the presence of the five transporter genes across eleven other endophyte genotypes which have also been characterised for their production of these three metabolites. Identifying corresponding patterns of secondary metabolite production with putative transporter presence could help with future characterisation of the transporter genes identified.

Table 3.10 Endophyte Information

Strain	Genetic Background	Alkaloids produced	gABC 4e	gABC 4c	gABC 4g	MFS E54	MFS E94
RS2	*E. festucae* x *E. typhina* x *E. baconii*	Ergovaline (H)	+	+	+	+	+
TF15 (AR555)	*E. festucae* x *E. baconii*	Peramine (H) Ergovaline Lolitrem	-	+	+	+	+
AR40	*N. lolii* (*E. festucae*)		+	+	+	+	+
FL1 (AR1501)	*E. festucae*	Lolitrem (H) Peramine (H) Ergovaline (H))	+	+	+	+	+
AR1	*N. lolii* (*E. festucae*)	Peramine	+	+	+	+	+
Lp19 (AR42)	*N. lolii* (*E. festucae*)	Ergovaline (H) Lolitrem (H) Peramine (H)	+	+	+	+	+
AR48	*N. lolii* (*E. festucae*)	None detected	+	+	+	+	+
AR66	*N. lolii* (*E. festucae*)	None detected	+	-	+	+	+
Lp2 (AR26)	*E. typhina* x *N. lolii* (*E. festucae*)	Ergovaline (L) Lolitrem Peramine	+	+	+	+	+
E8	*E. typhina*		-	+	+	-	-
AR501	*E. typhina* x *E. baconii*	Loline (L) Peramine (H)	-	-	+	-	+
Fp2 (AR1002)	*E.typhina* x *E. bromicola*	None detected	-	+	+	+	+

The endophyte name is shown as seen in Figure 3.23 as well as its AgResearch strain number (AR) if not already noted. The background states the endophyte genome, including the origin of any endophyte hybrids. Production of the alkaloids: peramine, ergovaline, loline and lolitrem are included and the levels seen, high (H) or low (L). The results of the strain distribution study are compared against this additional information

Table 3.10 shows the alkaloids each endophyte strain produces in its original background. *N. lolii* has the strongest expression of the alkaloids lolitrem, ergovaline and peramine which is why it was chosen to be used in this study. If there was a correlation between the alkaloids an endophyte produces and one of the five transporters studied it could be seen in Table 3.10. However the presence or absence of certain alkaloid compounds does not seem to be the factor affecting distribution of either the ABC or MFS transporter genes identified from *N. lolii* in this study. Additionally, none of the transporters identified in this study are clustered with characterised alkaloid gene clusters from endophyte such as peramine (Tanaka *et al.*, 2005), lolitrem (Young, 2005) , and ergovaline (Damien Fleetwood, personal communication).

The strain distribution study had also been undertaken with polyketide synthase (PKS) genes and non-ribosomal peptide synthase (NRPS) genes. These genes are commonly found in secondary metabolite genes clusters as their products are important in metabolite synthesis. The distribution pattern of the five identified ABC and MFS class of transporters was compared alongside the distribution of the NRPS and PKS genes in the same endophyte genotypes. A correlation in presence/absence of these gene types could suggest they are clustered in the same pathway. However no correlation could be seen between the transporter genes and the metabolite synthesis genes. The transportation of *N. lolii* secondary metabolites across the fungal cell membrane does not appear to be the role of these ABC and MFS transporter genes.

3.5 Sequencing and Annotation of Transporter Gene gABC 4e and Flanking DNA

The putative ABC gene gABC 4e was chosen to be sequenced. This gene was selected because it was an ABC gene which was identified using the degenerate PCR method and also was prominent in the EST database; early BLASTX results suggested that the gene fragment is a multidrug resistant protein similar to other fungal MDRs (Table 3.4). A λ library of Lp19 genomic DNA was screened using a DIG-labelled probe (Sections 2.5.2, 2.5.3, 2.7). The clone containing the gene was sequenced from an excised λ phagemid which was screened through two phage isolation rounds before the excision protocol was followed. The Lp19 DNA insert size was estimated to be approximately 6.5 kb from a *Bam* HI/*Eco* RI digest (data not shown).

3.5.1 Sequencing of a Lambda Phagemid Insert

The ABC 4e gene and flanking DNA was sequenced using Big Dye terminator biochemistry (Applied Biosystems) on an ABI 3100 Genetic Analyser with a 22cm capillary array (Section 2.3.8). In the first group of sequencing, M13 forward and reverse primers were used as well as the original ABC primer sets for gABC 4e, ABC FBE and gABC 5b (Table 2.1). New primers were designed approximately 100 bp upstream from the sequence ends, and contiguous sequences were joined using Vector NTI, ContigExpress program.

The complete fragment was sequenced in both orientations to ensure accuracy and complete coverage. The size of the insert sequenced was 6725 bp, which contained an ATP-binding region close to one end. The EST gene ABC FBE already comprised of the Walker A, B and ABC motifs and the genomic gABC 4e originated from degenerate primers nested specifically to these motifs at the protein's C-terminus.

3.5.2 Analysis of the Complete Sequence of an ABC Transporter Gene from *N. lolii*

The BLAST algorithm was used as a tool to confirm the sequence's identity as a putative ABC transporter and also to confirm predictions made as to the fragment of gene already sequenced form Section 3.1. BLASTX results indicated that the second

half of the gene was closest to the M13 reverse primer binding sites of the vector and was sequenced first.

The BLASTX results for part of the sequenced region (Table 3.11a) show excellent matches to known ABC transporter genes and many hypothetical transporters as well. The known genes that the gABC 4e query sequence aligns with includes multidrug protein transporters, including an MDR from *Emericella nidulans* (Andrade, et al., 2000); ABC4 from *Venturia inaequalis* (Schnabel, et al., 2001); and MDR1 from *Aspergillus fumigatus* (Tobin, et al., 1997).

Table 3.11 BLASTX Results of Insert gABC 4e's Predicted ORFs

3.11a **3866-5797 bp – ABC transporter multidrug resistance**

Accession	Protein Description	Score	e-value
XP365646	hypothetical protein *Magnaporthe grisea*	362	4e-98
BAE55073	unnamed protein product *Aspergillus oryzae*	362	4e-98
AAD43626	multidrug resistance protein MDR *Emericella nidulans*	358	3e-97
XP659904	hypothetical protein *Aspergillus nidulans*	358	3e-97
BAE64667	unnamed protein product *Aspergillus oryzae*	357	7e-97
AAB88658	multidrug resistance protein 1 *Aspergillus fumigatus*	357	1e-96
AAL57243	ABC multidrug transporter Mdr1 *Aspergillus fumigatus*	357	1e-96

3.11b **1372-2077 bp – Orotate phosphoribosyltransferase**

Accession	Protein Description	Score	e-value
AAC02431	orotate phosphoribosyl transferase *Metarhizium anisopliae*	315	1e-84
XP386340	orotate phosphoribosyltransferase *Gibberella zeae*	286	5e-76
AAB19948	orotate pyrophosphoribosyl transferase *Trichoderma reesei*	275	6e-73
XP962061	orotate phosphoribosyl transferase *Neurospora crassa*	269	6e-71
AAB24061	orotate phosphoribosyl transferase *Glomerella graminicola*	262	6e-69
EAQ88194	orotate phosphoribosyltransferase *Chaetomium globosum*	259	5e-68

Comparisons using BLASTX algorithms showed greatest similarity of the sequence to the second ATP-binding domain of known ABC transporters. This suggests that the original fragments were amplified from the COOH terminal ATP-binding region

(Figure 3.26) that the degenerate primers were designed to. The end of the gene appears to be 1 kb from the reverse end of the pBluescript SK- vector (Appendix III).

3.5.3 Searching for Exons: Gene Annotation

BLASTX results were used to identify sequence with good ABC transporter identity. However as ABC transporters are found in all species and carry out such a range of functions not all exons are well conserved. The BLASTX results (Table 3.11) identified the ABC transporter region as about 1.5 kb in length and positioned at 4228-5791 bp on the complete 6725 bp sequence. A full sized ABC transporter would contain two ATP-binding domains and cover approximately 4.5 kb of genomic sequence. Searching for similarity using the BLASTX algorithm on smaller sections of the phagemid insert identified two other hypothetical proteins flanking this putative ABC region. A putative orotate phosphoribosyltransferase on the reverse strand (Table 3.11b) is almost 2.5 kb from the identified ABC region between 1414-1755 bp (Figure 3.24); a small sequence of 341 bp was identified using the BLASTX algorithm. Another hypothetical protein may be at the M13 reverse end of the insert but the sequence is short and fragmentary and cannot be identified with any certainty.

To predict the intron and exon boundaries within the putative genes, two gene finder programs, SNAP and Genezilla were used. These genome annotations were completed by a Bioinformatition (Anar Khan, AgResearch New Zealand). SNAP (Korf, 2004) was run using a parameter model constructed from *Fusarium graminearum* predicted proteins (Stajich J., Duke University, http://fungal.genome.duke.edu/). Parameter models use a model species with a similar GC% to the target organism, in this case *N. lolii*, to find open reading frames. This is used when there is little information about the target species, and few genes of *N. lolii* have been completely sequenced so there is not enough annotated sequence available. The second program, Genezilla (Majoros, et al., 2004), uses the same *F. graminearum* model as mentioned above. Genezilla was formerly known as TigrScan and works in a similar way to SNAP: they just use different parameters.

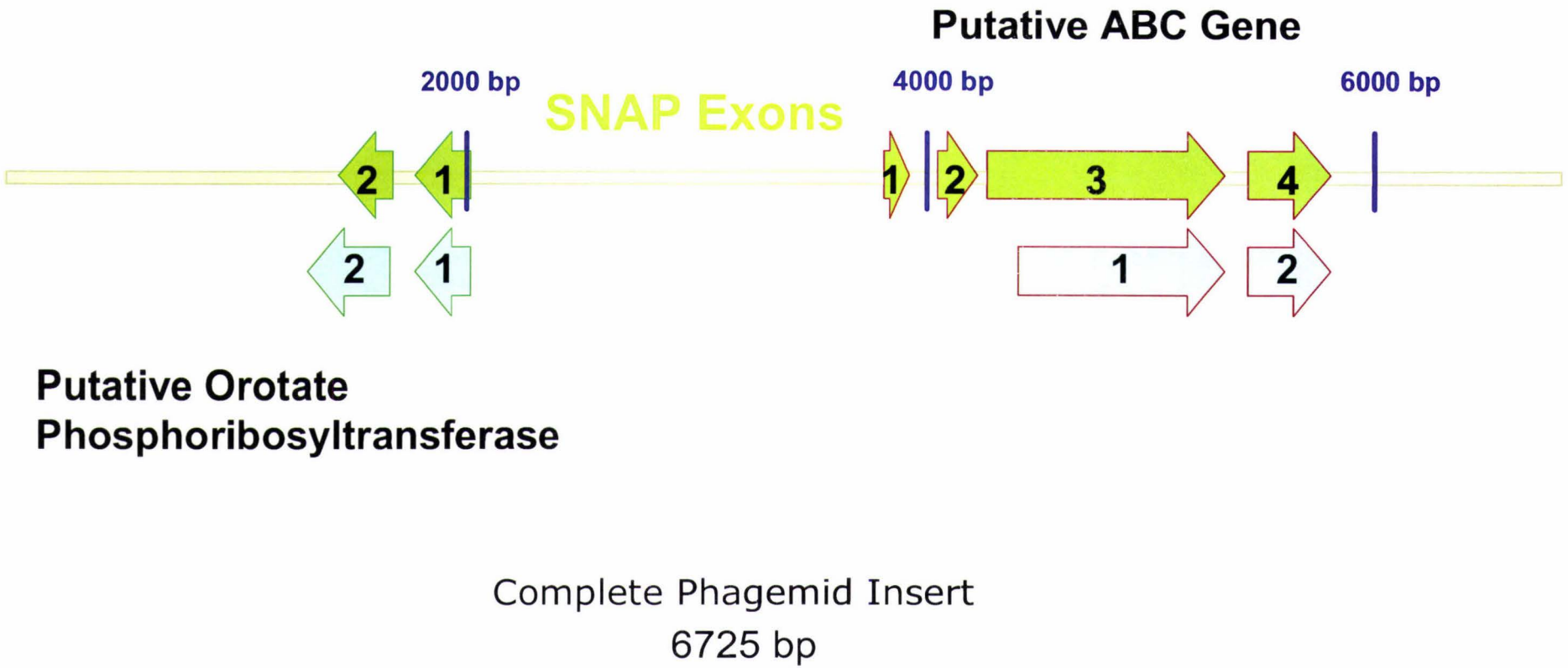
Both ORF prediction programs, SNAP and Genezilla, also predicted two genes on the 6.7 kb fragment submitted (Figure 3.24). An ORF on the sense strand corresponding to the ABC gene and an ORF on the reverse strand, for the putative orotate

phosphoribosyltransferase. The ABC gene has two exons according to Genezilla, exon 1 from 4447-5339 and exon two from 5440 -5797 bp (Table 3.12). However SNAP predicts four exons, with exons 3 and 4 seeming fairly certain as they correspond with the Genezilla prediction but it also predicts two smaller upstream ORFs. The start of SNAP exon 3 does not match with the beginning of Genezilla exon 1: BLASTX results from Table 3.11 overlapping the region of exons 3 and 4 as well as some ABC transporter gene similarity seen further upstream. The combined BLASTX and gene prediction results suggest that the exon predictions from SNAP are more accurate than the Genezilla results. The total length of all four of the SNAP predicted ORFs comes to 1552 bp, which is not long enough for a full sized 4.5 kb ABC transporter, but it may indicate that the gABC 4e gene is a half sized ABC transporter (2.5 kb) with one nucleotide binding region (NBR) and set of transmembrane regions (TMS) rather than two (Section 1.3.2).

Table 3.12 Predicted Open Reading Frames

Gene Predictor	Feature	Start	End	Strand
Genezilla	**mRNA**	**1372**	**2077**	**-**
	exon	1372	1727	-
	exon	1837	2077	-
Genezilla	**mRNA**	**4447**	**5797**	**+**
	exon	4447	5339	+
	exon	5440	5797	+
SNAP	**mRNA**	**995**	**2077**	**-**
	exon	995	1022	-
	exon	1505	1739	-
	exon	1837	2077	-
SNAP	**mRNA**	**3866**	**5797**	**+**
	exon	3866	3973	+
	exon	4097	4260	+
	exon	4314	5339	+
	exon	5440	5797	+

Figure 3.24 Phagemid Containing gABC 4e a Putative ABC Transporter Gene.



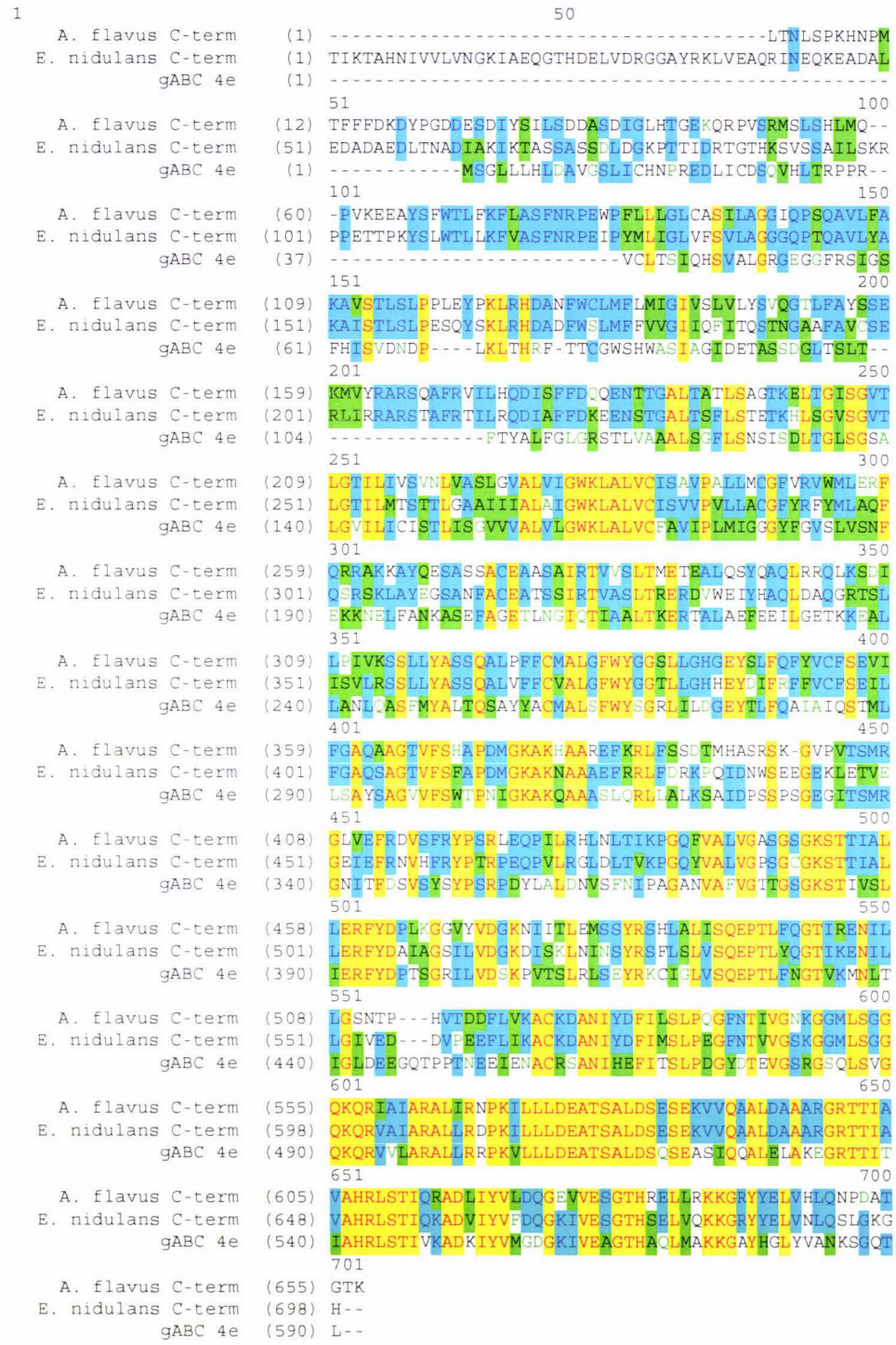
Contiguous sequences created by using a primer walking method to sequence the putative ABC gene gABC 4e were aligned using Clustal W (ContigExpress). The consensus sequence of the plasmid is represented here with the large arrows for the predicted exons and the orientation of their respective genes.

The putative Orotate phosphoribosyltransferase (ORPTase) gene upstream (2370 bp) from the putative ABC transporters (Figure 3.24) was located in the some position using the gene finder programs as had earlier been shown using BLASTX. The gene contains two exons 2077-1837, 1727-1372 bp as predicted by Genezilla and 2077-1837, 1739-1505 bp SNAP (Table 3.12). These exons are located on the antisense strand of the plasmid sequence with the gene divergently translated with respect to the putative ABC gene. As two genes were identified on the phagemid insert the *A. nidulans, F. graminarium*, *M. grisea*, *N. crassa* and *Ustilago maydis* genomes were checked for any similar ABC transporters of ORPTases that may also be similarly linked close to each other on these genomes (results not shown). However no synteny was seen on other fungal genomes for these two gene sequences.

Translation of the predicted ABC ORFs meant that the protein could be compared to known ABC transporters and the motifs aligned using Align X (Vector NTI). This protein alignment helped to confirm which ATP-binding domain had been sequenced (i.e. the N_2 or the COOH terminal NBR) and determine the orientation of the transmembrane domains and the ATP-binding domains enabling classification of gABC 4e into a defined class of the ABC gene family. The genes used in these alignments were *Aspergillus flavus* (AAB88655) and *Emericella nidulans* (AF071411) as the $[TMD_6\text{-}NBF]_2$ oriented ABCs, *Mycosphaerella graminicola* (CAB46280) and *Magnaporthe grisea* (EAA48214) represent the $[NBF\text{-}TMD_6]_2$ class of ABC transporters. The sequenced gene aligns best with the second half of the *A. flavus* and *E. nidulans* (Figure 3.25). The location of the 3' end of the gene and putative stop codon seem fairly certain as predicted by Genzilla and SNAP (Table 3.12). Figure 3.24 shows the Walker A, B and ABC motifs with the Walker B motif only 200 bp (60 aa) from the end of the gene, meaning that gABC 4e ends with an ATP-binding domain. The transmembrane region of this gene must come before the NBR so this putative ABC transporter is in a forward orientation TMD_6-NBR and therefore might belong to the MDR class of ABC transporters.

Alignments with Half sized transporter genes were also attempted (Results not shown). However Half-sized ABC transporters are more conserved than multidrug resistance ABC genes and these alignments did not find any regions of similarity of gABC 4e with known half-sized fungal ABC proteins.

Figure 3.25 Amino Acid Alignment of gABC 4e with the C-terminal of fungal ABCs



Translation of gABC 4e, putative introns removed, aligned with the C-terminus ends of ABC proteins from *A. flavus* (AAB88655) and *E. nidulans* (AF071411).

Neither the gene finder programs nor the BLASTX algorithm identified a second ATP-binding region within the putative ABC gene. It seemed likely that the initial NBR identified was the second of these regions, because of the similarity to the second half of ABC transporters using BLASTX alignments and because this was the ATP-binding region that the original degenerate primers had been designed to. Although some possible ABC motifs were seen upstream using smaller fragments in the BLASTX searches, the e-values were too high to be considered valid (Results not shown). The amino acids composing the N-terminal NBR of the ABC protein were expected to be less conserved, and so InterproScan (Zdobnov, 2001) a protein domain/motif searching program, was used to look for further NBR motifs.

InterproScan uses a collection of protein signature recognition programs including Pfam and other HMM algorithms to identify familiar protein motifs. Assessing the complete 6.7 kb of sequence with InterproScan found only the putative ABC and orotate phosphoribosyltransferase motifs on the ORFs that had been identified before. Screening only the region of sequence between 2100 and 3800 bp did not find any further motifs or protein domains. The putative ABC transporter identified may be a half sized transporter. Often two half sized ABC transporter products dimerise to form a functional 12 transmembrane spanning channel. A half sized ABC transporter would explain why only one set of ATP-binding motifs can be identified.

The exons identified by Genezilla were searched for intron splice sites, the splice positions can be confirmed by ensuring the translation of the product is kept in frame. The expected consensus sequences for *N. lolii* intron splice site were taken from Kupfer *et al.*, (2004) which had identified these recognition sites from five diverse fungal sequences (*S. cerevisiae*, *S. pombe*, *A. nidulans*, *N. crassa* and *C. neoformans*). The splice sites conserved between three species (excluding yeast species) are shown in Figure 3.26.

The splice sites were identified easily between SNAP predicted exons three and four for the putative ABC gene, this intron was also suggested by the program Genezilla and so seems fairly certain. The remaining splice sites are less certain. Removal of possible introns from between SNAP predicted exons one and two did not keep the product translated in the correct frame. Splice sites are not always clearly conserved in *N. lolii*

genes (Christine Voisey, personal communication), and few genes have been completely characterised to build a better model to predict them. Several attempts were made to gain a workable model for the exons and start site of gABC 4e by comparing BLASTX results with the exon predictions and the translated sequence alignments. Figure 3.27 includes the proposed splice sites as well showing three predicted exons. Predicted exon 1 (SNAP) shows a good translation initiation site consensus sequence (Figure 3.26, Figure 3.27) and may well be the start site for the gene if it is a half sized ABC transporter. SNAP exon 1 was extended from 108 bp to 351 bp, incorporating the first SNAP intron and some of exon 2 to better match the translated sequence alignments. The start of SNAP exon 3 (Exon 2 in figure 3.27) was moved upstream to include amino acids with good similarity from the BLASTX results. The intron/ exon boundary predictions still need to be experimentally confirmed; however Figure 3.27 shows the best predictions with the results already obtained.

Figure 3.26 Intron Splice Sites

Intron Splice/Branch Consensus Sites and Translation Initiation Sites

Consensus 5' donor sequence

```
      A
NG GU AGU
      G
```

Branch consensus

```
A  A C
 CU A
G  G T
```

Consensus 3' donor sequence

```
    C
UN AG
   T
```

Kupfer *et al.*, 2004

Acremonium chrysogenum Translation Initiation site consensus

```
  C             A A
C   N N C A     A T G G C
  A           C   C
```

Jekosch and Kück, 1999

The intron/exon junction consensus sequences shown are from the fungal species *A. nidulans*, *N. crassa* (Ascomycota) and *C. neoformans* (Basidiomycota). The translation initiation consensus is from *A. chrysogenum*. Conserved nucleotides are underlined, an N represents any of the four bases and bases that may be one residue or another are shown above and below the conserved line.

Figure 3.27 Complete gABC 4e Gene sequence.

```
                          M  S  G  L  L  L  H  L  D  A  V  G
    1  aagaaagcgc caaaaATGTC TGGCCTCTTG TTACACCTAG ATGCTGTAGG
          ·  S  L  I  C  H  N  P  R  E  D  L  I  C  D  S  Q  ·
   51  GAGCCTGATC TGTCACAACC CGAGAGAAGA TTTGATATGT GATAGCCAAG
       V  H  L  T  R  P  P  R  V  C  L  T  S  I  Q  H  S
  101  TGCATTTGAC CCGGCCACCC AGAGTTTGTT TGACCTCGAT CCAGCATTCC
        V  A  L  G  R  G  E  G  G  F  R  S  I  G  S  F  H
  151  GTCGCTTTAG GCCGCGGTGA GGGAGGGTTC AGGTCCATCG GAAGCTTCCA
       ·  I  S  V  D  N  D  P  L  K  L  T  H  R  F
  201  TATTTCGGTC GACAACGATC CCCTAAAGCT AACACACCGG TTCtagacaa

  251  cgtgcggctg gtcacattgg gcttctatag ctgggattga tgagactgca

  301  agcagcgatg gacttaccag tctgacattt acgtacgcct tgtttggact

  351  cggccgcagc actctgcaac ccagcccagt cgcagccgcc ggcctttgcc
                                  V  A  A  A  L  S  G  F  L·
  401  acgcatcacg gtaatgcagt aaGTAGCCGC CGCGTTATCT GGCTTCTTAT
       S  N  S  I  S  D  T  G  L  S  G  S  A  L  G  V
  451  CAAATTCTAT AAGCGACTTA ACTGGTCTGA GTGGCAGCGC CTTGGGCGTT
       I  L  I  C  I  S  T  L  I  S  G  V  V  V  A  L  V
  501  ATTCTTATCT GCATTTCCAC GCTTATATCA GGAGTAGTCG TTGCGCTCGT
       ·  L  G  W  K  L  A  L  V  C  F  A  V  I  P  L  M·
  551  CCTTGGTTGG AAACTGGCAC TCGTGTGCTT TGCCGTTATT CCTTTGATGA
       I  G  G  G  Y  F  G  V  S  L  V  S  N  F  E  K  K
  601  TCGGCGGAGG TTATTTTGGA GTTTCTCTTG TTAGCAACTT CGAGAAGAAG
       N  E  L  F  A  N  K  A  S  E  F  A  G  E  T  L  N
  651  AACGAACTAT TTGCTAATAA GGCCTCGGAG TTTGCTGGCG AGACGCTGAA
       ·  G  I  Q  T  I  A  A  L  T  K  E  R  T  A  L  A·
  701  TGGTATACAA ACTATCGCGG CGCTAACCAA GGAGCGGACC GCATTAGCCG
       E  F  E  E  I  L  G  E  T  K  K  E  A  L  L  A  N
  751  AATTCGAGGA AATACTCGGC GAGACTAAGA AGGAAGCTTT ACTCGCAAAC
       L  Q  A  S  F  M  Y  A  L  T  Q  S  A  Y  Y  A  C
  801  TTGCAAGCTT CCTTCATGTA TGCACTAACC CAGTCAGCCT ACTATGCCTG
       ·  M  A  L  S  F  W  Y  S  G  R  L  I  L  D  G  E  ·
  851  TATGGCGCTG AGCTTTTGGT ATAGCGGGCG GCTGATCCTT GACGGCGAGT
       Y  T  L  F  Q  A  I  A  I  Q  S  T  M  L  L  S  A
  901  ACACTTTGTT CCAAGCCATC GCTATCCAGT CAACCATGCT CCTGAGTGCC
        Y  S  A  G  V  V  F  S  W  T  P  N  I  G  K  A  K
  951  TACTCGGCTG GCGTTGTCTT CTCCTGGACT CCTAACATCG GAAAGGCTAA
       ·  Q  A  A  A  S  L  Q  R  L  L  A  L  K  S  A  I  ·
 1001  ACAGGCTGCT GCCTCACTAC AACGACTGCT AGCATTGAAA TCTGCAATCG
       D  P  S  S  P  S  G  E  G  I  T  S  M  R  G  N  I
 1051  ACCCCTCGTC TCCGTCTGGA GAGGGAATTA CCTCTATGCG CGGCAATATC
        T  F  D  S  V  S  Y  S  Y  P  S  R  P  D  Y  L  A
 1101  ACGTTCGACT CGGTTAGCTA TTCATACCCT TCCCGCCCGG ACTACCTGGC
       ·  L  D  N  V  S  F  N  I  P  A  G  A  N  V  A  F  ·
 1151  GTTGGACAAT GTCAGCTTCA ACATTCCAGC AGGCGCCAAT GTCGCCTTTG
       V  G  T  T  G  S  G  K  S  T  I  V  S  L  I  E  R
 1201  TTGGCACCAC AGGTTCTGGC AAGAGCACCA TCGTCTCGCT GATTGAGCGG
        F  Y  D  P  T  S  G  R  I  L  V  D  S  K  P  V  T
 1251  TTTTATGACC CAACCAGTGG GAGGATACTT GTGGACTCCA AACCTGTTAC
       ·  S  L  R  L  S  E  Y  R  K  C  I  G  L  V  S  Q·
 1301  GTCGCTCCGC TTGTCCGAAT ATCGAAAATG CATTGGTTTG GTCAGCCAAG
       E  P  T  L  F  N  G  T  V  K  M  N  L  T  I  G  L
 1351  AGCCCACCCT GTTCAACGGC ACCGTCAAGA TGAATTTGAC GATTGGTTTG
        D  E  E  G  Q  T  P  P  T  N  E  E  I  E  N  A  C
 1401  GACGAAGAGG GCCAAACACC ACCCACTAAT GAGGAAATCG AGAATGCGTG
       ·  R  S  A  N  I  H  E  F  I  T  S  L  P
 1451  TCGCTCAGCC AACATTCATG AATTCATTAC CAGCCTTCCG tgagtactca

 1501  aatttccatt atcttgtctt ttggtcacaa acaaaaacca tatcatcata
                                                  D  G  Y·
 1551  ccctctgcct acaaatcgct gacgaactct gttcgatagg GATGGCTACG
       D  T  E  V  G  S  R  G  S  Q  L  S  V  G  Q  K  Q
 1601  ACACGGAAGT TGGCAGTAGA GGCAGTCAGC TCAGTGTCGG ACAAAAGCAG
        R  V  V  L  A  R  A  L  L  R  R  P  K  V  L  L  L
```

```
1651    CGTGTGGTGC TCGCCCGGGC CCTGCTTCGG CGACCTAAGG TTCTGTTGCT
          D  E  A   T  S  A  L   D
1701    CGACGAGGCG ACGTCTGCGC TGGATTCGCA ATCAGAAGCC AGCATTCAAC

1751    AGGCGCTTGA GCTGGCTAAA GAGGGTAGGA CAACAATCAC GATTGCGCAT

1801    CGACTGAGTA CAATTGTAAA AGCGGACAAA ATTTACGTCA TGGGTGACGG

1851    CAAAATTGTC GAAGCGGGAA CACACGCTCA GTTGATGGCC AAGAAGGGTG

1901    CTTATCATGG ACTGTACGTG GCAAACAAAA GTGGGCAAAC TCTGtaa
```

Complete sequence of the gABC 4e gene sense strand. Translation of the exons is above the sequence and the Walker A, B and ABC motifs are in bold. Possible splice sites are underlined (5' donor, 3' donor and branch point).

Using nested degenerate primers an ABC gene gABC 4e was identified from the *N. lolii* Lp19 genome; this same gene was also included in the EST database. This gene showed expression both *in planta* and in cultured mycelia, with an interesting pattern of distribution seen across 12 diverse endophyte genotypes, and was chosen for complete sequencing for these reasons. Several methods were combined to determine the start stop codons and the intron/exon boundaries, with translation of the genes as determined so far, having 591 amino acids and one ATP-binding region identified. Further work will help to confirm the gABC 4e gene's true size and a possible function.

Chapter Four Discussion and Conclusions

4.1 Identifying ABC and MFS Genes in the *N. lolii* Genome

Degenerate PCR was used to amplify four unique ABC gene fragments from the genome of *N. lolii* strain Lp19. This method was used as there is not yet a complete sequence for this genome. The putative ABC transporters all showed significant similarity to toxin transporters from other fungal species such as ABC1 transporter (*Magnaporthe grisea*), Atr4 (*Mycosphaerella graminicola*) and BMR1 from *Botryotinia fuckeliana*.

ABC genes have a common ATP-binding motif that identifies them as belonging to the ABC transporter family. They are the ideal targets for degenerate primers and have been used to identify ABC transporters in other organisms. For *Mycosphaerella graminicola* (Stergiopoulos, 2002), degenerate PCR primers were designed to the conserved amino acid sequences for Walker A, Walker B and ABC signatures of the pleiotropic drug resistance (PDR) classes of ABCs, based on a phylogenetic analysis of yeast ABC proteins. Five genes were amplified that showed a high level of similarity to known fungal transporter genes from *Botrytis cinerea*, *Aspergillus nidulans* and others from *M. graminicola*. Similarly ABC genes were identified in *Aspergillus nidulans* by do Nasimento (2002), and in *A. fumigatus, A. flavus* by Tobin (1997) Both studies had alignments to a variety of eukaryotic ABC genes (MDR sub-class, Tobin *et al.*, 1997) to design their degenerate primers and then synthesised oligonucleotides to reflect the codon bias of *A. nidulans*. In this way they identified genes with similarity to drug transporters: *abcA-D* from *A. nidulans, AfuMDR1*, *AfuMDR2* from *A. fumigatus,* and *AflMDR1* (*A. flavus*).

For the *N. lolii* ABC experiment the first gene alignments were too diverse, containing transporter sequences from many fungal species and many sub-classes of ABCs. Decreasing the taxonomic diversity and the number of sequences included in the alignment, improved the design of degenerate primers sets. With *N. lolii,* nesting of the degenerate PCR products gave the best results in terms of both the specificity of the products amplified and the diversity of transporters recovered. The ABC gene fragments identified in this method came from 88 cloned PCR products, 72 of which showed similarity to known ABC transporter genes. However it is hard to be sure that the fragments encoded putative drug transporters on the basis of BLAST results because

drug transporters are the best characterised and the most prominent of the ABC sub classes in the public databases. With only 300 bp of sequence for each of these genes, inferring function is very difficult, especially when the regions amplified contain only the conserved motifs for binding ATP. Interestingly the gABC 4e fragment, which was later sequenced fully, should not have been amplified in the nested PCR using primers ABC 6F/7R (Figure 3.6). The fragment contained the Walker A motif which was upstream of the degenerate primer binding sites and yet fragments containing this sequence were represented 15 times out of 88 in the GA 300 transformations. When the complete gene was sequenced it was obvious the degenerate primers could not have bound in the region around the Walker B motif as the nucleotide sequences did not match, however it would appear that the more degenerate of the nucleotides, in particular the third codon position, have bound further upstream. The nucleotides are more conserved in the ATP-binding region than elsewhere and there must be enough similarity for binding of the primers in a location they were not intended for.

There were many duplicate copies of the same gene fragment in the ABC transporter library from the nested PCR experiment. An assembly of fragments using AlignX organised the sequences into five groups and two or more sequences were checked from each group using the BLAST algorithm. The gene fragments were narrowed down to four sequences: gABC 4a, 4c, 4e and 4g (Table 3.4), selected as possible ABC transporters and used in later experiments. The top BLAST hits of these fragments showed similarity to ABC transporter genes or hypothetical genes which contained the Walker A, B or ABC motifs. Sequence gABC 3f had close similarity with a hypothetical protein from *M. grisea*, a genome closely related to *N. lolii*. This uncharacterised hypothetical protein contained the ATP-binding motifs of an ABC transporter and was likely to be a transporter, but it was the only close match for the gABC 3F fragment.

Nesting of the primers also located three genes that were not ABC transporters at all, a hydoxy-amino acid/alcohol permease (HAAAP) and two elongation factor genes (Ef-3). These genes also contain an ATP-binding domain with the same Walker A, B and ABC motifs as seen in the ABC transporter genes, but they do not belong to the ABC transporter gene family. The degenerate primers were therefore working as hoped to identify ATP-binding domains. The number of ABC fragments recovered from the *N.*

lolii genome is consistent with other degenerate ABC primer studies on fungal species; five genes were isolated in *M. graminicola* (Stergiopoulos, 2002), and a total of four ABC transporter genes were recovered from *Aspergillus* species (do Nascimento *et al.*, 2002; Tobin *et al.*, 1997).

On the contrary, MFS genes were never amplified from *N. lolii* using degenerate PCR; it would likely be useful if only one or two very specific multidrug resistance MFS genes had been used as the basis for primer design. The original alignment was ambitious containing fungal MFS transporters with a variety of functions from a range of fungal species and this demonstrated the lack of conservation amongst MFS gene sequences. Attempts were made to narrow the primer degeneracy by aligning characterised MFS sequences; less degenerate primers were synthesised but the range of target sequences for their binding was also lessened and so no MFS transporters were cloned. Degenerate primers are not the way to identify several MFS transporters with a variety of functions.

From an *N. lolii* EST library a further four putative transporter genes (two ABC and two MFS) were identified by comparing their sequence similarity to that of transporter genes found in public databases. These EST sequences were between 500 and 800 bp and had the advantage of containing no introns. As with the genomic ABC sequences, BLASTX searches with these fragments showed similarity to fungal multidrug resistance transporters for both the ABC and MFS genes. Once again more sequence would be required before any assumptions could be made as to the function of these genes within the endophyte.

4.2 Expression of ABC and MFS Genes

Reverse transcriptase PCR experiments showed that all transporter genes tested were expressed in cultured *N. lolii* Lp19 cDNA. RT-PCR was undertaken to learn more about the roles of endophyte transporter genes in plant/endophyte symbiosis. Genes important in symbiosis may be up-regulated *in planta* compared to the cultured endophyte. The *Atr* genes ABC transporters of *A. nidulans* (Semighini, 2002) had their expression levels characterised against toxins using real time PCR. They found that the genes *AtrA-D* were induced by or up-regulated in their expression in, the presence of some antifungal compounds showing how these ABC transporters are involved in multidrug resistance.

These RT-PCR experiments are semi-quantitative, the levels of expression in each tissue may be compared for one gene but the amount of endophyte present in each tissue was not measured and must be kept in mind as a factor affecting the results. Also the expression levels can not be accurately compared over all the genes studied, however with these limitations in mind, some assumptions can be made as to the expression of the endophyte genes *in planta*.

No expression was seen of gABC 4a in either the blade tissues or the immature leaf tissues, and the expression of this gene in leaf sheaths, where endophyte concentration is highest, was very low. As gene expression *in planta* was an important criterion for continued research on a gene, gABC 4a was no longer considered a gene of interest.

The remaining genes were all expressed in plants, the strongest expression being in the sheath tissue. This difference in tissue expression was probably seen because sheath tissue contains a greater amount of mycelia than other tissues, with immature tissue having low levels of endophyte that are still growing and developing. *In planta* the metabolic activity of *Neotyphodium* endophyte was studied by Herd *et al.*, 1997. They found the highest concentration of endophyte metabolic activity was in the leaf sheath, with both mature leaf blades and the lower part of the emerging leaf (immature) showing 10 times less activity and the upper emerging leaf 100 times less than the leaf sheath. These results seem to be mimicked in the transporter gene expression for this study (Figure 3.20). The mature leaf tissues such as the blade and the sheath are the most concentrated for metabolite synthesis, and so high expression levels of the genes in these tissues, may have indicated the transporters' involvement in export of secondary metabolites or other roles in symbiosis. Therefore expression of the transporter genes was seen to be correlated with endophyte infection levels. The levels of expression for these transporters may remain fairly stable throughout the growth of the host.

The stability of expression seen in most of these genes suggests that they may be involved in an important function constitutive to the growth of the fungus; they may be sugar transporters or transport signalling molecules or other peptides around the fungus. If they were drug transporters the ABC genes are less likely to be the transporters of

secondary metabolites and more likely to be involved in transporting exogenous toxins such as plant defence compounds. Although none of the genes in this study appear to be highly up-regulated *in planta* compared to the cultured mycelia, this is not nessecarily initiative of the genes' functions and they may still play a role in drug transport.

MFS E54 shows less expression in culture than any of the ABC genes analysed in the project, and it was expressed at lesser levels in the leaf sheath and leaf blade tissues. MFS genes are often associated with gene clusters (Nierman *et al.*, 2005) and are more likely to be the carriers of secondary metabolites than ABC genes are. The alkaloids of interest in *N. lolii* lolitrem, ergovaline and peramine are produced from enzymes encoded by biosynthetic gene clusters, some of which have MFS genes nearby. MFS genes co-located with known biosynthetic gene clusters in filamentous fungi, such as the paxilline pathway from *Penicillium paxilli* have no homology with this MFS EST suggesting it is not involved in the export of any alkaloid metabolites studied so far.

MFS genes are often associated with gene clusters (Gardiner *et al.*, 2005; Schoonbeek *et al.*, 2003) and are more likely to be the carriers of secondary metabolites than ABC genes are. The alkaloids of interest in *N. lolii* lolitrem, ergovaline and peramine are produced from enzymes encoded by biosynthetic gene clusters, some of which have MFS genes nearby. MFS E54 shows less expression in culture, although it is seen faintly in leaf sheath and fainter in leaf blade tissues. This MFS may be up regulated *in planta* and therefore involved with symbiosis.

Another way to look at genes which are expressed differentially in the association between endophyte and grass is to do a suppressive subtractive hybridisation experiment (SSH); as was undertaken for the *N. coenophialum*- tall fescue symbiosis by Johnson *et al.*, 2004. SSH was used to find genes (both grass and endophyte) up-regulated or down-regulated in endophyte-infected plants, as well as in different tissue types and developmental stages. An up-regulated grass ABC gene was identified, similar to *Arabidopsis* and rice ABC transporters known to transport plant defence compound out of cells (Sanchez-Fernandez *et al.*, 2001; Xiong *et al.*, 2001). In *L. perenne* expression of this PDR5-like ABC transporter appeared similar in an endophyte-free plant compared with two different endophyte-infected grasses (Johnson

et al., 2004), and was thought to be involved in transporting plant compounds out of the grass cells or perhaps involved in transporting known or unknown fungal metabolites.

The SSH method of identifying genes of interest involved in the symbiotic association is complementary to microarray studies where large scale changes in groups of genes can be determined simultaneously. The ABC and MFS transporter genes identified in this experiment have been included in the design of a dual genome endophyte (*N. lolii* Lp19) and *L. perenne* Affymetrix gene chip. The expression pattern of these transporters will therefore be determined in future experiments where the conditions for induction of the genes will become clearer.

4.3 ABC and MFS Genes in the Genomes of Endophyte and Other Fungi

Protein domain analysis of the *Candida albicans* genome identified 27 ATP-binding regions (Braun *et al.*, 2005). Of these 13 appear to be full sized ABC transporters, six are half sized ABC transporters and the other ATP domains are found on non-transporter genes including elongation factor genes and amino acid permeases as was found for the *N. lolii* genomic ABC identification (Table 3.4). Yeast genomes also contain six half sized transporters (Gbelska *et al.*, 2006), these seem to play a role in intracellular transport rather than multidrug resistance.

MFS genes are far more abundant in fungal genomes (Braun et al., 2005), most likely because they tend to be more specific towards their substrates and ABC genes are able to transport a variety of substrates. *Neurospora crassa* has 39 ABC transporters, 115 MFS; *C. albicans* 71 MFS compared with the 19 ABC genes mentioned above. *Botrytis cinerea* has 46 ABC genes and *Magnaporthe grisea* has 175 identified MFS transporter genes (Braun et al., 2005; Yoder and Turgeon, 2001).

The BAC library of *N. lolii* (Lp19) proved to be a valuable tool in the resolving the relatedness of transporter gene fragments identified by degenerate PCR and by searching and EST database. Overall the pooled 96 well plate of BAC clones had 53 wells that contained one of the transporter genes screened for. This is an indication of just how frequent and therefore important ABC and MFS transporters are within the genome of *N. lolii*.

The BAC screening suggested that the two ABC ESTs were part of the same transporter genes that had already been identified: gABC 4g and gABC 4e isolated from *N. lolii* genomic DNA using the degenerate primer approach. Screening the BAC library indicated which of the gene fragments hybridise to the same clone. Clones gABC 4e and ABC FBE localised to the same BACs and when aligned were found to contain identical sequence. While the clones of gABC 4g and the EST D5A showed different nucleotide sequences by sequence comparison and alignment so it was interesting to find that primers for these gene fragments amplified products from the same BAC clone. The fragments are linked, so are likely to be a part of the same ABC gene or on separate genes but linked close to each other.

The BAC inserts average 140 kb of genomic sequence, it is possible that some inserts could include more than one transporter gene as ABCs are about 4.5 kb and MFS genes are less than half this size at approximately 1.8 kb. If only one transporter is seen in each of the positive wells, i.e. 53/1536 BAC clones contain one of the five transporter genes screened for, transporters are represented frequently within this BAC library. The *N. lolii* genome is approximately 30 Mb and may contain up to 15, 000 genes, with MFS transporters' representing 1% of the genes in the genome of yeast (*S. cerevisiae*) and between 1-3% in filamentous fungal genomes (*C. albicans* and *A. niger*) the indication is that the *N. lolii* genome will contain many more transporter genes of the ABC and MFS classes than the five identified in this study.

Another use of the BAC library screening was that it confirmed that some of the putative transporters being studied are unique. The primers designed to them do not overlap in the gene products that they amplify, Table 3.8 shows each primer pair amplified products from a unique selection of the BAC clones and so they are unlikely to hybridise to other unknown ABC genes in the *N. lolii* genome. This suggests that that a diversity of genes was recovered from the degenerative PCR experiments

Of the MFS genes that were gathered from the EST library only the primers for MFS E54 amplified products from the BAC library screening, the other MFS gene, MFS E94 was not represented in the BAC library, and unfortunately was not identified in screening of the *N. lolii* genomic DNA lambda library

Analysis of the *Aspergillus fumigatus* genome revealed 26 clusters of genes involved in secondary metabolite synthesis. ABC genes were found in 6 of these clusters and MFS genes were seen in 17. Some clusters contained more than one transporter, there were 27 MFS genes seen in the 16 clusters. Five of the clusters contain both ABC and MFS genes (Nierman *et al.*, 2005). This suggests that MFS genes are more likely to be associated with gene clusters than ABC genes. The ABC genes identified in this experiment are therefore less likely to be clustered with secondary metabolite genes, than the MFS genes found. There is no evidence from the BAC localisation or strain distribution studies that they cluster with known endophyte secondary metabolite genes, such as the ergovaline cluster (Damien Fleetwood, personal communication), the lolitrem gene clusters (Young *et al.*, 2005) or the peramine cluster (Tanaka *et al.*, 2005).

The strain distribution study looked at the presence of transporter genes isolated from *N. lolii* Lp19 in other strains and species of grass endophyte. The endophytes chosen for the strain distribution study were from the genera *Epichloë* (*Neotyphodium* is the asexual anamorph of this genus) or *Neotyphodium*, which are hosted by several pasture grass species. Production of secondary metabolites, such as bioactive alkaloids, can differ between endophyte species. *N. lolii* for example is able to produce the insecticides ergovaline and lolitrem B (that causes rye grass staggers in livestock) and the feeding deterrent peramine. It can not produce the alkaloid lolines: these are made by the strain AR501 which infects tall fescue.

The strains for the study were selected across a range of diverse endophytes that were compared for relatedness to each other and assembled into a phylogenetic tree (Section 3.4.2) The *Neotyphodium* endophytes: *N. lolii* (Lp19, AR40, AR48, AR66), *N. coenophialum* (RS2) *and N. uncinatum* (Fp2) are found in the grass hosts perennial ryegrass, tall fescue and meadow fescue respectively and make a combination of lolitrem, peramine, loline and ergovaline alkaloids. Other *Epichloë* endophyte species *E. typhina* (E8) and *E. festucae* (Fl1) have a broad host range including *Festuca*, *Lolium*, and *Koeleria spp* and are able to produce a similar range of the characterised secondary metabolites. Endophyte biochemistry is highly diverse, and further uncharacterised bioactive secondary metabolites are also expected to be produced these strains.

The putative *N. lolii* ABC gene, gABC 4e (this gene was fully sequenced, see Section 3.5) was not present in TF 15, E8, AR501 and Fp2. These are all strains with an *E. typhina* endophyte background. Only one gene was present in all strains tested the ABC gene ABC D5A (gABC 4g), and may well be a constitutively expressed transporter, although further strains would have to be tested to confirm this. gABC 4c was not found in the only genotype that is seen to produce the alkaloid loline (AR 501) and was also missing from AR 66 a strain of *N. lolii*. The MFS genes were present in E8 (an *E. typhina* strain) and MFS E54 was also missing from AR 501. The reliability of this screen could be ensured by testing with a second set of primers. This would determine that any negative results are due to the absence of the transporter gene in the endophyte genome and not the result of a change in the nucleotide sequence of the species or strain.

The distribution patterns of these five putative transporters were compared with the presence or absence of endophyte non-ribosomyl peptide synthatase (NRPS) genes and polyketide synthase (PKS) genes in the same twelve genotypes. These two gene classes are important for the biosynthesis of secondary metabolites and are major components of fungal gene clusters. Unfortunately none of the gene distributions matched, suggesting perhaps that none of the five genes screened correlate with these genes, some of which are likely to be involved in secondary metabolite production.

4.4 Gene Analysis

The DNA insert which contains the gABC 4e gene is 6.7 kb, and contains at least one other gene, a putative orotate phosphoribosyltransferase gene (in the reverse orientation) as well as the putative ABC gene. A small fragment of another hypothetical protein may be present downstream of the ABC; however it is not possible to deduce the function of this gene using homology searches.

Gene finder programs SNAP and Genezilla were used to identify the possible open reading frames in the sequence and the BLASTX results helped to find splice recognition sites. The putative ABC gene has similarity to the C-terminal end of ABC proteins from *Magnaporthe grisea* and *Aspergillus fumigatus* that are characterised as multidrug transporters. The last two exons identified by both gene finder programs

corresponded almost exactly to each other and seem fairly reliable with respect to the end of the gene, with both programs identifying the same stop codon. ATP-binding motifs compromising the Walker A, B and ABC motifs were recognisable in one frame of the translation and are located in the final two exons consistent with ABC genes of the MDR class in a forward orienntation.

The start ATG codon of the gene was not definitively identified using bioinformatics tools. It was not possible to ascertain whether the gene was a half sized ABC transporter or a full size transporter as there was no match to ABC transporters in the region that would comprise the N-terminal half of the transporter. Half sized transporters contain only one set of six transmembrane spanning structures and one ATP-binding region, these may be in the forward (TMD_6-NBF) or the reverse topology (NBF-TMD_6). Protein dimerisation of half sized transporters is thought to occur before functioning as a channel through the cellular membrane. Half sized transporters are true orthologues; six identical genes have been found conserved in all yeasts and filamentous fungi. These ABC genes have been protected from the loss and duplication that is common among the full sized transporters and they seem to be involved not in multidrug resistance but rather in intracellular roles (Gbelska *et al.*, 2006). In the genome of the yeast *S. cerevisiae*, Gbelska *et al.*, (2006) classified the six half sized ABC transporters as involved in functions with peroxisomal fatty acyl CoA transport (*Pxa1, Pxa2*), mitochondrial transport of peptides (*Mdl1, Mdl2*), heavy metal transport (*Atm1*) and a transporter located in the endoplasmic reticulum (*Adp1p*).

BLASTX alignments of gABC 4e DNA showed strong homology with full sized transporters. No half sized transporters were recovered from BLASTX results. Alignments of this gene with half sized transporter from other fungi such as *S. cerevisiae* and *A. fumigatus* indicated no substantial regions of conservation. Thus particularly in the light of observations by Gbelska *et al.*, (2006) that there is substantial evidence of true orthologues for half sized transporters in fungi, this suggests that gABC 4e may be a full length transporter. If this is the case, there is room to accommodate a full length gene of approximately 4.5 kb before the start of the next open reading frame (Orotate phosphoribosyl transferase). However attempts to find transmembrane domains and NBRs in DNA upstream of the conserved ABC sequence

were not successful, as any domains identified had a very low e-value suggesting they were not valid motifs.

The gABC 4e is in the forward orientation whether it is a half or a full sized ABC gene, with the NBF (ATP-binding region) at the C-terminal of the gene. With this topology gABC 4e is likely to belong to the MDR class of ABC transporters and would be performing the role of a drug transporter of some kind. Since the degenerate primers used to amplify this gene were originally designed to transporters of this class, discovery of a MDR ABC transporter would not be unexpected. It was not possible within the time frame of this project to confirm the intron/ exon boundaries or indeed the transcriptional start site to determine whether the gene is a full length or a half size transporter, or if the gene is a pseudogene. Further work to confirm this would be valuble.

The orotate phosphoribosyl transferase (OPRTase) gene upstream of the ABC gene is in the reverse orientation in comparison with the ABC gene, and is a protein involved in the biosynthesis of pyrimidine nucleotides. This protein performs the fifth in a six step biosynthetic pathway towards synthesis of pyrimidines (Nara *et al.*, 2000). The ORPTase gene does not appear to be linked with an ABC gene in any publicly available fungal genomes, nor do they seem part of a gene cluster. No synteny was observed for these gene sequences in the *A. nidulans*, *F. graminarium*, *M. grisea*, *N. crassa* or *Ustilago maydis* genomes.

4.5 Conclusion

The diversity of sequence at the nucleotide level means that degenerate primers that are able to amplify several classes of ABC genes are difficult to design and need to be nested for greater specificity. However ABC genes were identified from the *N. lolii* genome using this technique, and 4 or 5 different genes were isolated, consistent with other fungal degenerate PCR studies (Nascimento *et al.*, 2002; Stergiopoulos *et al.*, 2002).

Sequence diversity of MFS genes means that degenerate PCR is not a useful tool to isolate these genes. There are many more MFS than ABC genes to be found. In fungal genomes these genes are frequently associated with clusters, in the genome of *A.*

nidulans two MFS genes were often found in one cluster (Nierman *et al.*, 2005). To identify more MFS genes perhaps the homologues of specific genes need to be located; designing specific primers and assessing them against a gradient of less stringent annealing temperatures to increase the chances of amplifying the required product.

As expected *N. lolii* does contain expressed ABC and MFS genes. Based on searches using the BLASTX algorithm the six identified (two MFS, four ABC) are the most similar to multidrug transporters, and may function in this way. So far no similarity has been identified with other classes of transporter genes or even half sized ABC genes; further sequence is required to develop this hypothesis.

The six transporter genes are all expressed in cultured endophyte, and most are expressed *in planta* as well, suggesting they are involved in the transport of substrates in most conditions. *In planta* expression is the highest in the sheath tissue, the most mature plant tissue, which also has the highest concentration of endophyte.

By screening the genes across a selection of endophyte strains, some interesting patterns emerge with gABC 4e (the gene completely sequenced) which is not present in any strains that colonise fescue grasses as their host and ABC D5A as the only gene present in all strains selected for amplification. This suggests that gABC 4e is not involved in "house keeping" activities often associated with half size transporters

BLASTX homology and ORF prediction programs both suggest that ABC transporter gABC 4e is the size of a half sized transporter, of approximately 2 kb, however half sized transporters in fungi are fairly well characterised as intracellular transporters and no similarity is seen with these genes from other fungi using the BLASTX algorithm. Any ABC motifs upstream are very poorly conserved as even the identified ABC gene ORF was not greatly similar to ABC genes in the known database beyond the ATP-binding motifs. Further work is required to identify this genes' start codon, and to confirm the intron/exon junctions. Rapid amplification of cDNA ends (RACE) of *N. lolii* cDNA would be especially useful for confirming the start and finish of the sequence.

Sequencing of the gABC 4e gene shows the ABC transporter is located adjacent to a pyrimidine synthesis gene. The ABC gene although proximal to the orotate phosphoribosyltransferase gene is probably not involved in the export of pyrimidines, as no synteny was seen with other fungal genomes. BLASTX results continue to show similarity of gABC 4e with multidrug ABC genes but not at a low enough e-value to infer gene function.

The isolation of five transporter genes from an endophyte genome is similar to what has been found in other ABC degenerate PCR studies, this gives an indication of the validity of the technique for isolating ABC genes, but further ABC and MFS genes remain to be identified in the *N. lolii* genome. Completing the sequence of these genes will improve the chances of allocating a putative function based on sequence similarity. The inclusion of these ABC and MFS transporter genes onto a gene chip will enable them to be analysed in many experiments to come, and may indicate the conditions that these genes and other similar ABC and MFS genes are induced in.

References

Al-Samarrai, T.H., and Schmid, J. (2000). A simple method for extraction of fungal genomic DNA. *Letters in Applied Microbiology* **30:** 53-56.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215:** 403-410.

Andrade, A. C., Del Sorbo, G., Van Nistelrooy, J. G., and Waard, M. A. (2000a). The ABC transporter AtrB from *Aspergillus nidulans* mediates resistance to all major classes of fungicides and some natural toxic compounds. *Microbiology* **146 (Pt 8):** 1987-1997.

Andrade, A. C., Van Nistelrooy, J. G. M., Peery, R. B., Skatrud, P. L., and De Waard, M. A. (2000b). The role of ABC transporters from *Aspergillus nidulans* in protection against cytotoxic agents and in antibiotic production. *Molecular and General Genetics* **263:** 966-977.

Bauer, B. E., Wolfger, H., and Kuchler, K. (1999). Inventory and function of yeast ABC proteins: about sex, stress, pleiotropic drug and heavy metal resistance. *Biochimica et Biophysica Acta* **1461:** 217-236.

Borges-Walmsley, M. I., and Walmsley, A. R. (2001). The structure and function of drug pumps. *Trends in Microbiology* **9:** 71-79.

Braun, B. R., van Het Hoog, M., d'Enfert, C., Martchenko, M., Dungan, J., Kuo, A., Inglis, D. O., Uhl, M. A., Hogues, H., Berriman, M., *et al.* (2005). A human-curated annotation of the *Candida albicans genome. PLoS Genetics* **1:** 36-57.

Broccardo, C., Luciani, M., and Chimini, G. (1999). The ABCA subclass of mammalian transporters. *Biochimica et Biophysica Acta* **1461:** 395-404.

Bullock, W.O., Fernandes, J.M., and Short, J.M. (1987). *XL-1* Blue: a high efficiency plasmid transforming *recA Escherichia coli* strain with β-galactosidas seclection. *Biotechniques* **5:** 376-378.

Callahan, T. M., Rose, M. S., Meade, M. J., Ehrenshaft, M., and Upchurch, R. G. (1999). CFP, the putative cercosporin transporter of *Cercospora kikuchii*, is required for wild type cercosporin production, resistance, and virulence on soybean. *Molecular Plant-Microbe Interactions* **12:** 901-910.

Chang, P. K., Yu, J. J., and Yu, J. H. (2004). aflT, a MFS transporter-encoding gene located in the aflatoxin gene cluster, does not have a significant role in aflatoxin secretion. *Fungal Genetics and Biology* **41:** 911-920.

Christensen, M. J. (1995). Variation In The Ability Of *Acremonium* Endophytes Of *Lolium-Perenne*, *Festuca-Arundinacea* And *F-Pratensis* To Form Compatible Associations In The 3 Grasses. *Mycological Research* **99:** 466-470.

Christensen MJ, Bennett RJ, Bryan G, Goetz, C, Johnson R, Simpson W, Tapper B, Voisey C. (2004). Is colonization of leaves by *Epichloe/ Neotyphodium* made possible by non-tip hyphal extension? In: Roberts C; West CP ed. 5th International Symposium on *Neotyphodium*/ Grass Interactions. 203.

Christensen, M.J., Leuchtmann, A., Rowan, D.D., and Tapper, B.A. (1993). Taxonomy of *Acremonium* endophytes of tall fescue (*Festuca arundinacea*), medow fescue (*F. pratensis*) and perennial ryegrass (*Lolium perenne*). *Mycological Research* **97:** 1083-1092.

Clay, K., and Schardl, C. (2002). Evolutionary origins and ecological consequences of endophyte symbiosis with grasses. *The American Naturalist* **160:** S99-S127.

De Waard, M. A. (1997). Significance of ABC Transporters in Fungicide Sensitivity and Resistance. *Pesticide Science* **51:** 271-275.

Del Sorbo, G., Schoonbeek, H., and De Waard, M. A. (2000). Fungal transporters involved in efflux of natural toxic compounds and fungicides. *Fungal Genetics and Biology* **30:** 1-15.

Faeth, S. H., and Sullivan, T. J. (2003). Mutualistic asexual endophytes in a native grass are usually parasitic. *The American Naturalist* **161:** 310-325.

Fleissner, A., Sopalla, C., and Weltring, K. M. (2002). An ATP-binding cassette multidrug-resistance transporter is necessary for tolerance of *Gibberella pulicaris* to phytoalexins and virulence on potato tubers. *Molecular Plant-Microbe Interactions* **15:**102-108.

Fletcher, L. R., and Harvey, I. C. (1981). An association of a *Lolium* endophyte with ryegrass staggers. *New Zealand Veterinary Journal* **29:** 185-186.

Friijter, A.C.J., Zhang, Z., van Damme, M., Wang, G.-L., Ronald P.C., and Michelmore, R.W. (1997). Construction of a bacterial artifical chromosome library containing large *Eco*RI and *Hind*III genomic fragments of lettuce. *Theoretical and Applied Genetics* **94:** 390-399.

Gardiner, D. M., Cozijnsen, A. J., Wilson, L. M., Pedras, M. S., and Howlett, B. J. (2004). The sirodesmin biosynthetic gene cluster of the plant pathogenic fungus *Leptosphaeria maculans*. *Molecular Microbiology* **53:** 1307-1318.

Gardiner, D. M., Jarvis, R. S., and Howlett, B. J. (2005). The ABC transporter gene in the sirodesmin biosynthetic gene cluster of *Leptosphaeria maculans* is not essential for sirodesmin production but facilitates self-protection. *Fungal Genetics and Biology* **42:** 257-263.

Gbelska, Y., Krijger, J., and Breunig, K. D., (2006). Evolution of gene families: the multidrug resistance transporter genes in five related yeast species. *FEMS Yeast Research* **6:** 345-355.

Ginn, S. L., Brown, M. B., and Skurray, R.A. (2000). The TetA(K) tetracycline/H+ Antiporter from *Staphylococcus aureus*: mutagenesis and functional analysis of motif C. *Journal of Bacteriology* **182:** 1492-1498.

Haarmann, T., Machado, C., Lubbe, Y., Correia, T., Schardl, C. L., Panaccione, D. G., and Tudzynski, P. (2005). The ergot alkaloid gene cluster in *Claviceps purpurea*: Extension of the cluster sequence and intra species evolution. *Phytochemistry* **66:** 1312-1320.

Hayashi, K., Schoonbeek, H. J., and De Waard, M. A. (2002a). Bcmfs1, a novel major facilitator superfamily transporter from *Botrytis cinerea*, provides tolerance towards the

natural toxic compounds camptothecin and cercosporin and towards fungicides. *Applied and Environmental Microbiology* **68:** 4996-5004.

Hayashi, K., Schoonbeek, H. J., and De Waard, M. A. (2002b). Expression of the ABC transporter BcatrD from *Botrytis cinerea* reduces sensitivity to sterol demethylation inhibitor fungicides. *Pesticide Biochemistry and Physiology* **73:** 110-121.

Herd, S., Christensen, M. J., Saunders, K., Scott, D. B. and Schmid, J. (1997). Quantitative assessment of *in planta* distribution of metabolic activity and gene expression of an endophytic fungus. *Microbiology* **143:** 267-275.

Higgins, C. F. (1992). ABC Transporters: From Microorganisms to Man. *Annual Review of Cell Biology* **8:** 67-113.

Johnson, L. J., Johnson, R. D., Schardl, C., and Panaccione, D. G. (2004). Identification of differentially expressed genes in the mutualisic association of tall fescue with *Neotyphodium coenophialum*. *Physiological and Molecular Plant Pathology* **63:** 305-317.

Keller, N. P., and Hohn, T. M. (1997). Metabolic pathway gene clusters in filamentous fungi. *Fungal Genetics And Biology* **21:** 17-29.

Korf, I. (2004). Gene finding in Novel genomes. *BMC Bioinformatics* **5:** 59

Kunkel, B. A., Grewal, P. S., and Quigley, M. F. (2004). A mechanism of acquired resistance against an entomopathogenic nematode by *Agrotis ipsilon* feeding on perennial ryegrass harboring a fungal endophyte. *Biological Control* **29:** 100-108.

Kupfer, D. M., Drabenstot, S. D., Buchanan, K. L., Lai, H., Zhu, H., Dyer, D.W., Roe, B. A. and Murphy, J. W. (2004) Introns and splicing elements of five diverse fungi. *Eukaryotic Cell.* **3:** 1088-1100.

Leuchtmann, A. (1994). Isozyme relationships of *Acremonium endophytes* from twelve *Festuca* species. Mycological Research 98: 25-33.

Majoros, W. H., Pertea, M., and Salzberg, S. L. (2004). TigrScan and GlimmerHMM: two open source *ab initio* eukaryotic gene finders. *Bioinfrmatics* **20(16):** 2878-2879.

Malinowski, D. P., and Belesky, D. P. (2000). Adaptations of Endophyte-Infected Cool-Season Grasses to Environmental Stresses. *Crop Science* **40:** 923-940.

Martin, J. F., Casqueiro, J., and Liras, P. (2005). Secretion systems for secondary metabolites: how producer cells send out messages of intercellular communication. *Current Opinion in Microbiology* **8:** 282-293.

Martinoia, E., Klein, M., Geisler, M., Bovet, L., Forestier, C., Kolukisaoglu, U., Muller-Rober, B., and Schulz, B. (2002). Multifunctionality of plant ABC transporters--more than just detoxifiers. *Planta* **214:** 345-355.

Moon, C. D., Miles, C. O., Jarlfors, U., and Schardl, C. L. (2002). The evolutionary origins of three new *Neotyphodium* endophyte species from grasses indigenous to the Southern Hemisphere. *Mycologia* **94:** 694-711.

Nara, T., Hshimoto, T., and Aoki, T. (2000). Evolutionary implications of the mosaic pyrimidine-biosynthetic pathway in eukaryotes. *Gene* **257:** 209-222.

do Nascimento, A. M., Goldman, M. H., and Goldman, G. H. (2002). Molecular characterization of ABC transporter-encoding genes in *Aspergillus nidulans*. *Genetics and Molecular Research* **1:** 337-349.

Nierman, W. C., Pain, A., Anderson, M.J., Wortman, J. R., Kim, H. S., Arroyo, J., *et al.* (2005). Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. *Nature* **438:** 1151-1156.

Panaccione, D. G., Johnson, R. D., Wang, J., Young, C. A., Damrongkool, P., Scott, B., and Schardl, C. L. (2001). Elimination of ergovaline from a grass-*Neotyphodium* endophyte symbiosis by genetic modification of the endophyte. *Proceedings of the National Acadamy of Sciences U S A* **98:** 12820-12825.

Panaccione, D.G., Tapper, B.A., Lane, G.A., Davies, E., and Fraser, K. (2003). Biochemical outcome of blocking the ergot alkaloid pathway of a grass endophyte. *Journal of Agriculture and Food Chemistry* **51:** 6429-6437.

Pao, S. S., Paulsen, I. T., and Saier, M. H., Jr. (1998). Major facilitator superfamily. *Microbiology and Molecular Biology Reviews* **62:** 1-34.

Paulsen, I. T. (2003). Multidrug efflux pumps and resistance: regulation and evolution. *Current Opinion in Microbiology* **6:** 446-451.

Pitkin, J. W., Panaccione, D. G., and Walton, J. D. (1996). A putative cyclic peptide efflux pump encoded by the *TOXA* gene of the plant-pathogenic fungus *Cochliobolus carbonum*. *Microbiology* **142 (Pt 6):** 1557-1565.

Rowan, D.D. (1993). Lolitrems, peramine and paxilline: mycotoxins of the ryegrass/ endophyte interaction. *Agriculture, Ecosystems and Environment* **44:** 103-122.

Saier, M. H., Jr., and Paulsen, I. T. (2001). Phylogeny of multidrug transporters. *Seminars in Cell & Developmental Biology* **12:** 205-213.

Sanchez-Fernandez, R., Davies, T. G., Coleman, J. O., and Rea, P. A. (2001). The *Arabidopsis thaliana* ABC protein superfamily, a complete inventory. *Journal of Biological Chemistry* **276:** 30231-30244.

Schardl, C. L., Leuchtmann, A., and Spiering, M. J. (2004). Symbioses of grasses with seedborne fungal endophytes. *Annual Review of Plant Biology* **55:** 315-340.

Schardl, C.L., Leuchtmann, A., Tsai, H., Collet, M.A., and Scott, D.B. (1994). Origin of a fungal symbiont of perennial ryegrass by interspecific hybridisation of a mutualist with the ryegrass choke pathogen, *Epichloë typhina*. *Genetics* **136:** 1307-1317.

Schnabel, G., Schnabel, E., and Jones, A. L. (2001). Genes encoding multidrug resistance like proteins in *Venturia inaequalis*. *Modern Fungicides and Antifungal Compounds* **3:** 297-305.

Schoonbeek, H. J., van Nistelrooy, J. G. M., and de Waard, M. A. (2003). Functional analysis of ABC transporter genes from *Botrytis cinerea* identifies BcatrB as a transporter of eugenol. *European Journal of Plant Pathology* **109:** 1003-1011.

Semighini, C. P., Marins, M., Goldman, M. H., and Goldman, G. H. (2002). Quantitative analysis of the relative transcript levels of ABC transporter *Atr* genes in *Aspergillus nidulans* by real-time reverse transcription-PCR assay. *Applied and Environmental Microbiology* **68:** 1351-1357.

Spiering, M. J., Lane, G. A., Christensen, M. J., and Schmid, J. (2005a). Distribution of the fungal endophyte *Neotyphodium lolii* is not a major determinant of the distribution of fungal alkaloids in *Lolium perenne* plants. *Phytochemistry* **66:** 195-202.

Spiering, M. J., Moon, C. D., Wilkinson, H. H., and Schardl, C. L. (2005b). Gene clusters for insecticidal loline alkaloids in the grass-endophytic fungus *Neotyphodium uncinatum*. *Genetics* **169:** 1403-1414.

Spiering, M. J., Wilkinson, H. H., Blankenship, J. D., and Schardl, C. L. (2002). Expressed sequence tags and genes associated with loline alkaloid expression by the fungal endophyte *Neotyphodium uncinatum*. *Fungal Genetics And Biology* **36:** 242-254.

Stergiopoulos, I., Gielkens, M. M., Goodall, S. D., Venema, K., and De Waard, M. A. (2002a). Molecular cloning and characterisation of three new ATP-binding cassette transporter genes from the wheat pathogen *Mycosphaerella graminicola*. *Gene* **289:** 141-149.

Stergiopoulos, I., Zwiers, L. H., and De Waard, M. A. (2002b). Secretion of natural and synthetic toxic compounds from filamentous fungi by membrane transporters of the ATP-binding cassette and major facilitator superfamily. *European Journal of Plant Pathology* **108:** 719-734.

Stergiopoulos, I., Zwiers, L. H., and De Waard, M. A. (2003). The ABC transporter MgAtr4 is a virulence factor of *Mycosphaerella graminicola* that affects colonization of substomatal cavities in wheat leaves. *Molecular Plant-Microbe Interactions* **16:** 689-698.

Sullivan, T. J., and Faeth, S. H. (2004). Gene flow in the endophyte *Neotyphodium* and implications for coevolution with *Festuca arizonica*. *Molecular Ecology* **13:** 649-656.

Tanaka, A., Tapper, B.A., Popay, A., Parker, E.J., and Scott, B. (2005). A symbiosis expressed non-ribosomal peptide synthetase from a mutualistic fungal endophyte of perennial ryegrass confers protection to the symbiotum from insect herbivory. *Molecular Microbiology* **57:** 1036-1050.

Tobin,M.B., Peery,R.B. and Skatrud,P.L. (1997). Genes encoding multiple drug resistance-like proteins in *Aspergillus fumigatus* and *Aspergillus flavus*. *Gene* **200:** 11-23.

Tutulan-Cunita, A. C., Mikoshi, M., Mizunuma, M., Hirata, D., and Miyakawa, T. (2005). Mutational analysis of the yeast multidrug resistance ABC transporter Pdr5p with altered drug specificity. *Genes to Cells* **10:** 409-420.

Upchurch, R. G., Rose, M. S., and Eweida, M. (2001). Over-expression of the cercosporin facilitator protein, CFP, in Cercospora kikuchii up-regulates production and secretion of cercosporin. *FEMS Microbiology Letters* **204:** 89-93.

Upchurch, R. G., Rose, M. S., Eweida, M., and Callahan, T. M. (2002). Transgenic assessment of CFP-mediated cercosporin export and resistance in a cercosporin-sensitive fungus. *Current Genetics* **41:** 25-30.

Urban, M., Bhargava, T., and Hamer, J. E. (1999). An ATP-driven efflux pump is a novel pathogenicity factor in rice blast disease. *The Embo Journal* **18:** 512-521.

van den Brule, S., and Smart, C. C. (2002). The plant PDR family of ABC transporters. *Planta* **216:** 95-106.

Vermeulen, T., Schoonbeek, H., and De Waard, M. A. (2001). The ABC transporter BcatrB from Botrytis cinerea is a determinant of the activity of the phenylpyrrole fungicide fludioxonil. *Pest Management Science* **57:** 393-402.

Voss, T., Schulte, J., and Tudzynski, B. (2001). A new MFS-transporter gene next to the gibberellin biosynthesis gene cluster of *Gibberella fujikuroi* is not involved in gibberellin secretion. *Current Genetics* **39:** 377-383.

Walton, J. D. (2000). Horizontal gene transfer and the evolution of secondary metabolite gene clusters in fungi: an hypothesis. *Fungal Genetics And Biology* **30:** 167-171.

Xiong, L., Lee, M. W., Qi, M., and Yang, Y. (2001). Identification of defense-related rice genes by suppression subtractive hybridization and differential screening. *Molecular Plant-Microbe Interactions* **14:** 685-692.

Yoder, O. C., and Turgeon, B. G. (2001). Fungal genomics and pathogenicity. *Current Opinion in Microbiology* **4:** 315-321.

Young, C.A. (2005) The indole-diterpene gene cluster from the ryegrass endophyte, *Neotyphodium lolii*, is required for the biosynthesis of Lolitrem B, a biopretective alkaloid. *PhD thesis, Massey University New Zealand.*

Young, C., McMillan, L., Telfer, E., and Scott, B. (2001). Molecular cloning and genetic analysis of an indole-diterpene gene cluster from *Penicillium paxilli*. *Molecular Microbiology* **39:** 754-764.

Young, C. A., Bryant, M. K., Christensen, M. J., Tapper, B. A., Bryan, G. T., and Scott, B. (2005). Molecular cloning and genetic analysis of a symbiosis-expressed gene cluster for *lolitrem biosynthesis* from a mutualistic endophyte of perennial ryegrass. *Molecular Genetics And Genomics* **274:** 13-29.

Zdobnov E. M. and Apweiler R. (2001). InterProScan - an integration platform for the signature-recognition methods in InterPro *Bioinformatics* **17(9):** 847-8.

Zwiers, L. H., and De Waard, M. A. (2000). Characterization of the ABC transporter genes *MgAtr1* and *MgAtr2* from the wheat pathogen *Mycosphaerella graminicola* Quantitative analysis of the relative transcript levels of ABC transporter *Atr* genes in *Aspergillus nidulans* by real-time reverse transcription-PCR assay. *Fungal Genetics And Biology* **30:** 115-125.

Zwiers, L. H., Stergiopoulos, I., Gielkens, M. M., Goodall, S. D., and De Waard, M. A. (2003). ABC transporters of the wheat pathogen *Mycosphaerella graminicola* function as protectants against biotic and xenobiotic toxic compounds. *Molecular Genetics and Genomics* **269:** 499-507.

APPENDICES

Appendix I Media

All media were prepared with Milli-Q water and autoclaved at 121 °C for 15 minutes to sterilize.

For selective media supplements were added to final concentrations of: Ampicillin (100 μg/ml), Kanamycin sulphate (50 μg/ml), 5-bromo-4-chloro-3-indolyl-b-D-galactoside in dimethylformamide (X-gal) (60 μg/ml), Chloramphenicol (25 μg/ml), Tetracycline (15μg/m).

A1.1 *E.coli*

LB (Luria-Broth) Broth (g/L): Tryptone 10.0, Yeast extract 5.0, NaCl 5.0 (pH 7.0)
LB (Luria-Broth) Agar (g/L): Tryptone 10.0, Yeast extract 5.0, NaCl 5.0, Agar 15.0
0.7% Top agarose (g/L): Tryptone 10.0, NaCl 5.0, Agarose 7.0

A1.2 *N. lolii*

PD (Potato Dextrose) Broth (g/L): Potato dextrose broth Potato 24.0
PD (Potato Dextrose) Agar (g/L): Potato dextrose broth 24.0, Agar 15.0

Appendix II Common Buffers and Solutions

TAE buffer: 40 mM Tris acetate, 11.4 ml/L glacial acetic acid, 1 mM EDTA (pH 8.5).
SOC medium: 2% Tryptone, 0.5% Yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose.

A2.1 Genomic DNA Isolation Buffers

A2.1.1 Protoplasting Method

OM buffer: 1.4 M $MgSO_4$, 100 mM stock Na_2HPO_4 (pH 5.8)

STC buffer: 1M sorbitol, 50mM Tris-HCl (pH 8.0), 50mM $CaCl_2$

A2.1.2 High Molecular Weight Method

1x lysis buffer: 400 mM Tris acetate, 200 mM Sodium Acetate, 10 mM EDTA, 10% (w/v) SDS

Polysaccharide precipitator: 5 M NaCl

Phenol Solution: Phenol,1% 8-hydroxyquinoline, 2 M NaOH, 1 M Tris-HCl

A2.2 Southern Blotting, DIG Hybridisation Reagents

Denaturation Buffer: 0.4 M NaOH, 1.5 M NaCl

Neutralisation Buffer: 0.3 Tris.HCl (pH 7.2), 1M NaCl

20x SSC: 3 M NaCl, 0.3 M sodium citrate, pH 7.0

Wash Solution I: 2 x SSC, 0.1% (v/v) SDS

Wash Solution II: 0.5 x SSC, 0.1% (v/v) SDS

DIG Maleic Acid Buffer: 0.1 M Maleic Acid, 0.15 M NaCl, pH 7.5

DIG Tween Washing Buffer: 0.1 M Maleic acid, 0.15 M NaCl, (pH 7.5); 0.3% (v/v) Tween 20

Blocking solution: 1% (v/v) blocking reagent (Roche Applied Science), in Maleic acid buffer (see above)

Antibody solution: Anti-Digoxigenin AP diluted 1:10000 in blocking solution

Detection buffer: 100 mM Tris-HCl pH 9.5, 100 mM NaCl, 50 mM $MgCl_2$,

DIG Probe stripping buffer: 0.2 M NaOH, 0.1% (w/v) SDS

2x SSC: 0.3 M NaCl, 0.03 M sodium citrate pH 7.0

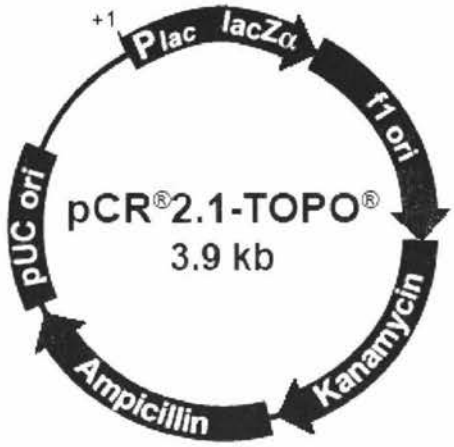
A2.3 Reagents for Library Screening

SM Buffer 5.8 g NaCl, 2 g $MgSO_4.7H_20$, 50 ml 1M Tris-HCl (pH 7.5), 2% (w/v) gelatine/ L

Appendix III Vector Maps

A 3.1 pCR 2.1-TOPO cloning vector (Invitrogen)

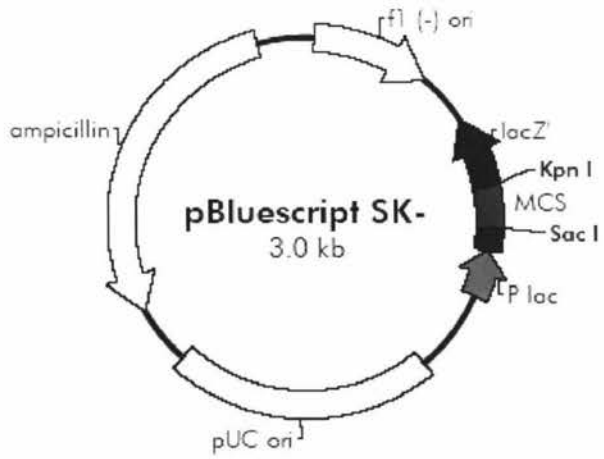
Contains M13 forward and reverse priming sites, multiple cloning sites, Kanamycin and ampicillin antibiotic resistance, T7 promoter binding site.



A3.2 pBluescript SK(-) (Stratagene)

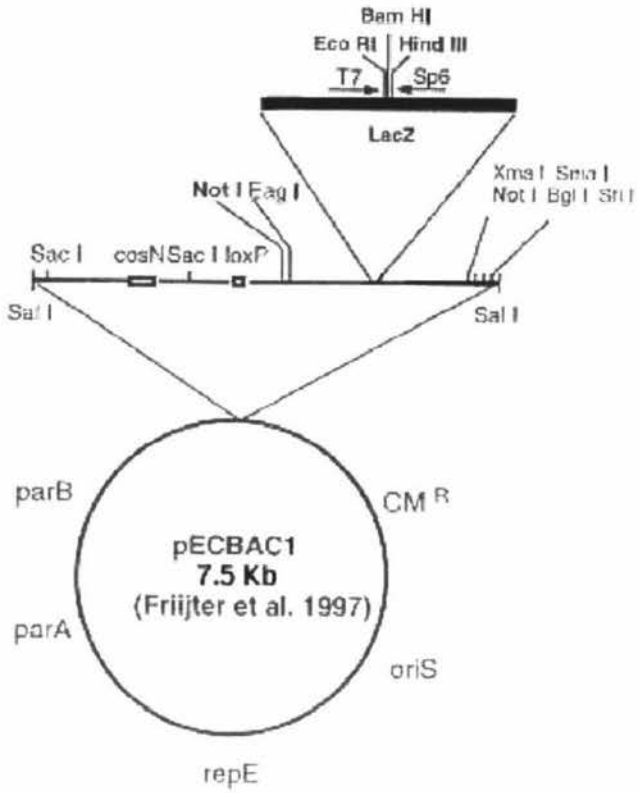
Has M13 forward and reverse primer binding sites, a multiple cloning site and Ampicillin resistance.

pBluescript® SK(–) Vector Map



A3.3 pECBAC1 (Amplicon Express)

Contains Chloramphenicol resistance.



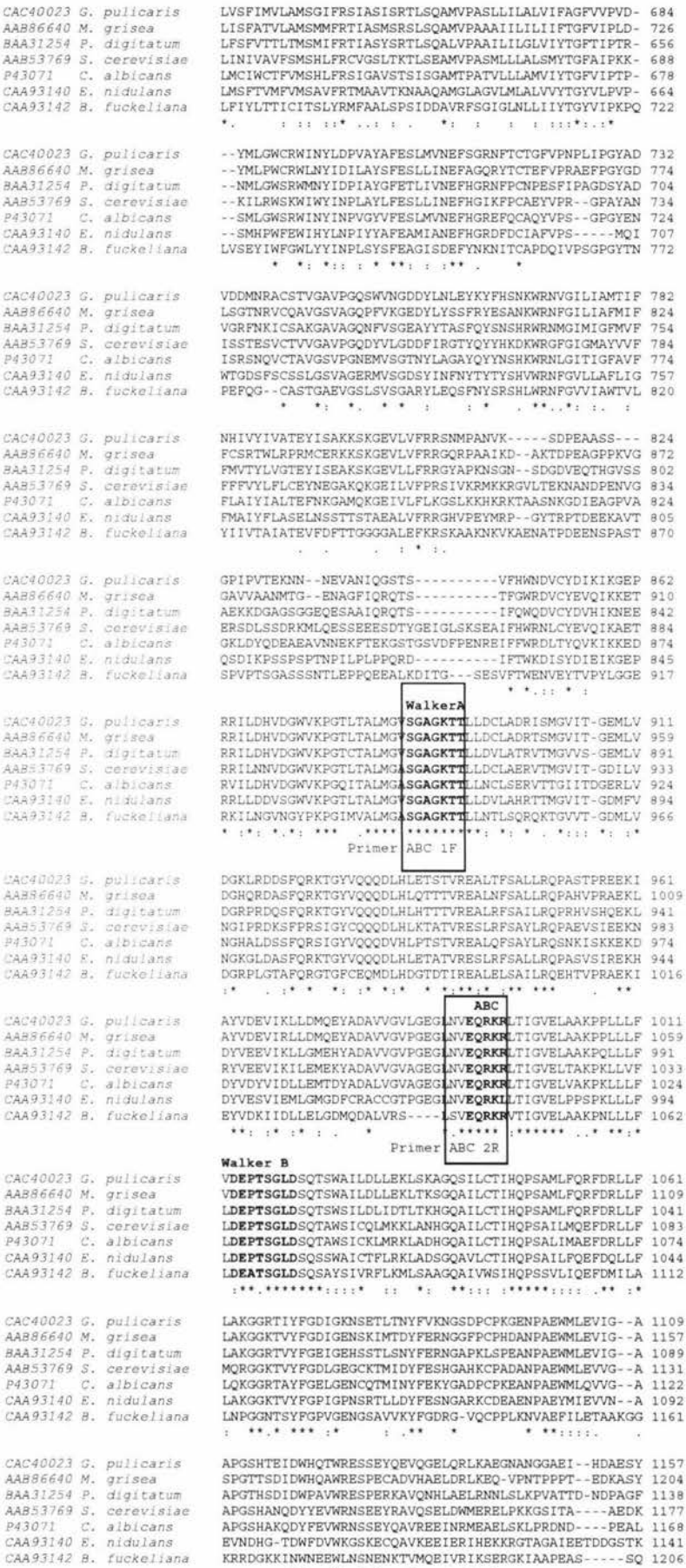
Appendix IV Additional Transporter Alignments

A4.1 ABC Alignment 1

Clustal W alignment of seven diverse ABC transporter genes. The motifs (Walker A, B and ABC) are indicated in bold and the primers are included in boxes.

```
CAC40023 G. pulicaris     -------------------------MATPDAN-----MSSTRSEQSSHD 19
AAB86640 M. grisea        MSQPVEDPSHDQARNDNAQTTTDTGNASMPKTNGHDQESSATGISSSPAD 50
BAA31254 P. digitatum     -------------------------------------------MEARRHD 7
AAB53769 S. cerevisiae    -------------------------MPEAKLNNNVNDVTSYSSASSSTE 24
P43071   C. albicans      -------------------------MSDSKMSSQDESKLEKAISQDSSS 24
CAA93140 E. nidulans      -------------------------MGVPDELPPGSSETDTIVSSSQPT 24
CAA93142 B. fuckeliana    ----MFYTHAERMYAERPIEGREGAQIPLHSDNYRGEPIQPHDSSSSESE 46

CAC40023 G. pulicaris     TIVNNELSTN-------EKPLQSAPAPGDQTSSTDDEDDGPQTEEMVRRH 62
AAB86640 M. grisea        TLMDKEKQAAPTPEPSTEKAAVPDPAASVASDTAEDEFFDDNDSEQRRRN 100
BAA31254 P. digitatum     PTAS----------------VNTEDTANEKSEVGDEKYTDAEVTR----- 36
AAB53769 S. cerevisiae    NAAD----------------LHNYNGFDEHTEARIQKLARTLTAQSMQNS 58
P43071   C. albicans      ENHS----------------INEYHGFDAHTSENIQNLARTFTHDSFKDD 58
CAA93140 E. nidulans      NRSP----------------MDLISEAESLNLRRIATNQSKAQCRPG 55
CAA93142 B. fuckeliana    INNDGTWGEHE--------QGGPVNCSTAMHEYEDLRHQLTHLSKTRSQK 88

CAC40023 G. pulicaris     SIVRDLARNYTNTS------HHFNGSNADLFNAADP--ASPLNPSSENFN 104
AAB86640 M. grisea        EMVQQLARTYTSRSNASAAADEYGNANPFLIASEDP--DSPLNPSGNNFK 148
BAA31254 P. digitatum     -LAQQLTRQSTRFS-----VSPQNAENPFIETHED----STLNPHSGNFK 76
AAB53769 S. cerevisiae    TQSAPNKSDAQSIF-----SSGVEGVNPIFSDPEAPGYDPKLDPNSENFS 103
P43071   C. albicans      SSAGLLKYLTH--------MSEVPGVNPYEHEEIN---NDQLNPDSENFN 97
CAA93140 E. nidulans      SAAVP----------------SHDNPPNDDLED----ATLDPNSASFS 83
```

```
CAA93142 B. fuckeliana    SVATARKASAMKRTMTNASRRSRATAGEDIEAQEEEEKLEEAAADEDEFV 138
                              .                .   .*


CAC40023 G. pulicaris     ARAWAKAMSKSMNENGAGF--RQSGLCFQDMN-VFGYGAETDYQKDVGNV 151
AAB86640 M. grisea        AYAWAKAIAGMVAAEGGSF--RTIGICFQNMN-VFGFGAATDFQKTVSNV 195
BAA31254 P. digitatum     AKNWMKNLLAIQSRDPERYPKRQAGLAFKNLS-VHGFGSPTDYQKDVANS 125
AAB53769 S. cerevisiae    SAAWVKNMAHLSAADPDFYKPYSLGCAWKNLS-ASGASADVAYQSTVVNI 152
P43071   C. albicans      AKFWVKNLRKLFESDPEYYKPSKLGIGYRNLR-AYGVANDSDYQPTVTNA 146
CAA93140 E. nidulans      LEKWLR--AAVSDASQQGLSTPSGGIPLQKSDRPRAQTAALQLQPTVGSV 131
CAA93142 B. fuckeliana    LGDFLKD--GHFEKRNTKGSAKKVGVIYKNLT-VQGVGATSTYVKTLPSA 185
                             : :                 *   :.    .          : .


CAC40023 G. pulicaris     WL---------GLPSMAQQLFKPNAGKRRIDILRGFDGVVNAGEMLVVLG 192
AAB86640 M. grisea        WL---------EAANMLRTAVG-MGKTTRIDILRGFNGVVRNGEMLVVLG 235
BAA31254 P. digitatum     VL---------EIGAFFRTMAG--TGKQKIQILRDFDGLVKNGEMLIVLG 164
AAB53769 S. cerevisiae    PY---------KILKSGLRKFQRSKETNTFQILKPMDGCLNPGELLVVLG 193
P43071   C. albicans      LW---------KLATEGFRHFQKDDDSRYFDILKSMDAIMRPGELTVVLG 187
CAA93140 E. nidulans      LT---------APLRS--RVCCDTGEYEPRRILHGFDGVMKTGELLLVLG 170
CAA93142 B. fuckeliana    IMGTFGPDLYKLLTRFIPALPKPGSNGQRRDLIHDFTGCVRDGEMLLVLG 235
                                                  ::: : . :. **: :***
                          Walker A
CAC40023 G. pulicaris     PPGSGCSTFLKSISGETNGIYVDDSTYFNYNGIAADEMHKHHKGETIYTA 242
AAB86640 M. grisea        PPGSGCSTFLKTIAGETNGLNVDQSAYFNYQGLSAEEMHKRHRGEAIYTA 285
BAA31254 P. digitatum     RPGSGCSTFLKTIAGEMNGIFKDANSHMNYQGISDKEMRNQFRGEAIYTA 214
AAB53769 S. cerevisiae    RPGSGCTTLLKSISSNTHGFDLGADTKISYSGYSGDDIKKHFRGEVVYNA 243
P43071   C. albicans      RPGAGCSTLLKTIAVNTYGFHIGKESQITYDGLSPHDIERHYRGDVIYSA 237
CAA93140 E. nidulans      RPGTGCSTFLKAVCGETNGLHIDADSVLHYNGVSQQRMMKEFKGEVVYNQ 220
CAA93142 B. fuckeliana    RPGSGCSTFLKAIS-NKRGDYAGVTGEVSYGGISAEEQAKHYRGEVNYNE 284
                           **:**:*:**::. :  *   .    . * * : .   ...:*:. *.


CAC40023 G. pulicaris     EVDVHFPMLSVGDTLTFAARARCPQNLPSGINHNQYSEHMRDVVMAMYGI 292
AAB86640 M. grisea        EVDVHFPQLSVGDTLTFAANARAPRRAPPGVSKTLFANHIRDVVMAIFGI 335
BAA31254 P. digitatum     ETDVHFPQLSVGNTLKFAALARAPRNRLPGVSRDQYAEHMRDVVMAMLGL 264
AAB53769 S. cerevisiae    EADVHLPHLTVFETLVTVARLKTPQNRIKGVDRESYANHLAEVAMATYGL 293
P43071   C. albicans      ETDVHFPHLSVGDTLEFAARLRTPQNRGEGIDRETYAKHMASVYMATYGL 287
CAA93140 E. nidulans      EVDKHFPHLTVRQTLEFGAAARTPAHRFQNMSRDEFASYAASVVMAIFGL 270
CAA93142 B. fuckeliana    EDDQHFPSLTVEQTLDFSLLNKTKKHEKGDIP------TIITALLKMFGI 328
                          * * *:* *:* :**      :   .  .:         . :   *:
                                          ABC                 Walker B
CAC40023 G. pulicaris     SHTVNTQVGDNYIRGVSGGERKRVTIAEATLSNAPFQCWDNSTRGLDSAN 342
AAB86640 M. grisea        SHTINTRVGNEYIRGVSGGERKRVTIAEAALSGAPLQCWDNSTRGLDSAN 385
BAA31254 P. digitatum     SHTINTRVGNDFIRGVSGGERKRVSIAEATLCGSPLQCWDNSTRGLDSAN 314
AAB53769 S. cerevisiae    SHTRNTKVGNDIVRGVSGGERKRVSIAEVSICGSKFQCWDNATRGLDSAT 343
P43071   C. albicans      SHTRNTNVGNDFVRGVSGGERKRVSIAEASLSGANIQCWDNATRGLDSAT 337
CAA93140 E. nidulans      SHTHNTKVGNDFVRGVSGGERKRVSIAEMALAMTPFAAWDNSSRGLDSAT 320
CAA93142 B. fuckeliana    SHTRHTLVGDAFVRGVSGGERKRVSIAETLPTKKTVVSWDNSTRGLDASP 378
                          *** :* **:  :**********:***       . .***::****::


CAC40023 G. pulicaris     AIEFCKTLRLQSELFGQTCAVSIYQAPQTAYDLFDKALVIYEGRQIFFGP 392
AAB86640 M. grisea        AIEFCKTLRVCTRLFQTTACVSIYQAPQSAYDMFDKAVVLYEGYQIYFGP 435
BAA31254 P. digitatum     ALEFCKTLNLMSKYSGTTCAVAIYQASQSAYDVFDKVTVLYEGRQIYFGR 364
AAB53769 S. cerevisiae    ALEFIRALKTQADISNTSATVAIYQCSQDAYDLFNKVCVLDDGYQIYYGP 393
P43071   C. albicans      ALEFIRALKTSAVILDTTPLIAIYQCSQDAYDLFDKVVVLYEGYQIFFGK 387
CAA93140 E. nidulans      ALKFVQALRLSADLAGAAHAVAIYQASQSIYEVFDKVTVLYEGRMIFFGP 370
CAA93142 B. fuckeliana    ALDYANSLRVMTDISNRTTLVTLYQAGEQIYELMDKVLVIDEGRMVYSGP 428
                          *:.: .:*.  :      :  :::**. :  *::::*. *: :*  :: *


CAC40023 G. pulicaris     ADEAKAYFIGLGFECPDRQTTPDFLTSMTAPSERVVRPGFENK---VPRT 439
AAB86640 M. grisea        ADEAKQYFVNLGFECPAR-TTPDFLTSMTAPHERIVRPGFEGK---APRT 481
BAA31254 P. digitatum     TTEAREFFTNMGFHCPDRQTTADFLTSLTSPAERVVKPGFEKM---VPRT 411
AAB53769 S. cerevisiae    ADKAKKYFEDMGYVCPSRQTTADFLTSVTSPSERTLNKDMLKKGIHIPQT 443
P43071   C. albicans      ATKAKEYFEKMGWKCPQRQTTADFLTSLTNPAEREPLPGYEDK---VPRT 434
CAA93140 E. nidulans      TGTAKEYFERMGWVCPARQTTGDFLTSITNPLERKARAGMEDV---VPKT 417
CAA93142 B. fuckeliana    AGDAKKYFERIGYYCPPRKTTADFLTACTDPAERRFQKDFEGP---IPKG 475
                          :  *: :*  :*: ** * ** ****: * * **    .        *:


CAC40023 G. pulicaris     PDEFATCWKQSQQYQIVRAEVESYKSLYPIN--GSSADAFRENKQSAQAK 487
AAB86640 M. grisea        PEEFAIAWENSAEYTALQADIEEYKSSHPIN--GPDAEAFRKSRAAQQGR 529
BAA31254 P. digitatum     PDEFAKGWKNSAAYKELQKEIDDYNTQYPIG--GESFQQFVESRKAMQSK 459
AAB53769 S. cerevisiae    PKEMNDYWVKSPNYKELMKEVDQRLLNDD----EASREAIKEAHIAKQSK 489
P43071   C. albicans      AQEFETYWKNSPEYAELTKEIDEYFVECER---SNTRETYRESHVAKQSN 481
CAA93140 E. nidulans      PKDFEIYWRQSPEYKTLLGEMTGVRDQHPTGNDEQASAELRARKENSQSR 467
CAA93142 B. fuckeliana    PVELEKAFRESEDYRLLQEDVESYERMLHET-DHADARQFKASVEETKSK 524
                          . ::   : :*  *  :  ::                        :..


CAC40023 G. pulicaris     GQRLKSPFTLSYMQQVQLCLWRGFKRLVGSPGVTIFQLIANTVVALIASS 537
AAB86640 M. grisea        GQRPNSPYTLSFYQQTKLCLWRGWKRLLGDPTLTVGALFANTLMALVISS 579
BAA31254 P. digitatum     GQRAKSPYTLSVAEQVQICVTRGFQRLKSDYSLTISALIGNTIMALIVGS 509
AAB53769 S. cerevisiae    RARPSSPYTVSYMMQVKYLLIRNMWRLRNNIGFTLFMILGNCSMALILGS 539
P43071   C. albicans      NTRPASPYTVSFFMQVRYGVARNFLRMKGDPSIPIFSVFGQLVMGLILSS 531
CAA93140 E. nidulans      NSRAASPYILSIPMQIKLNTKRAYQRIWNDMSSTMSTVVGQIVIALITGV 517
CAA93142 B. fuckeliana    TVGPRSPYTVSFFRQVLACTKREVELTLGDKTTLYTKFFIIISNSLIVGS 574
                               **: :*   *       *       ..       ..     .*: .


CAC40023 G. pulicaris     LFYNMEPT--TGDFFKRGAVLFLAVLSNAFASALEILTQYAQRPIVEKHA 585
AAB86640 M. grisea        IFFNLQMT--TSSFFQRGALLFFACLLNGFAAALEILILFAQRPIVEKHD 627
BAA31254 P. digitatum     VFYQLPDD--VTSFYSRGALLFFAVLLNSFSSALEILTLYAQRPIVEKQA 557
AAB53769 S. cerevisiae    MFFKIMKKGDTSTFYFRGSAMFFAILFNAFSSLLEIFSLYEARPITEKHR 589
P43071   C. albicans      VFYNLSQT--TGSFYYRGAAMFFAVLFNAFSSLLEIMSLFEARPIVEKHK 579
CAA93140 E. nidulans      VFYDSPNT--TAGFQSKGGTLFYAVLLNALTAMSEITSLYSQRPIVEKQA 565
CAA93142 B. fuckeliana    LFHGQSSN--TLGNFSRGGTLFFSILFLGWLQLSELMKAVGGRPIISRHK 622
                          :*.       .     :*. :* : * .     *:      *** .::


CAC40023 G. pulicaris     RYAFHHASAESFASILVDMPYKITNSILFNLTLYFMTNLNRDAGAFFFYL 635
AAB86640 M. grisea        RYALYHPSAEAVASMLCDMPYKVFNAIVFNLVLYFMANLRREPGAFFFYL 677
BAA31254 P. digitatum     RYAMYHPFAEAISSMLCDMPYKILNAITFNVTLYFMTGLRQNAGAFFTFM 607
AAB53769 S. cerevisiae    TYSLYHPSADAFASVLSEIPSKLIIAVCFNIIFYFLVDFRRNGGVFFFYL 639
P43071   C. albicans      KYALYRPSADALASIISELPVKLAMSMSFNFVFYFMVNFRRNPGRFFFYW 629
CAA93140 E. nidulans      SYAFYHPATEAIAGVVSDVPVKFLLAVAFNVIMYFLANLRREPAQFFIYF 615
CAA93142 B. fuckeliana    DYAFYRPSAVVVARVVQDFPLLLLQVVPFSIVVYFITGLDVDAGKYFIYV 672
                           *:::::. :  .: :: :.*  .   : *.. .**:..:  : . :* :
```


WalkerA
Primer ABC 1F
ABC
Primer ABC 2R
Walker B

CAC40023 *G. pulicaris*, ABC Transporter; AAB86640 *M. grisea*, ABC1; BAA31254 *P. digitatum*, PMR1; AAB53769 *S. cerevisiae*, ABC transporter; P43071 *C. albicans*, CDR1; CAA93140 *E. nidulans*, ATP-binding drug transporter; CAA93142 *B. fuckeliana*, ATP-binding multidrug transporter

A4.2 ABC Alignment 3

Clustal W alignment of three ABC transporter genes of the MDR class. The motifs (Walker A, B and ABC) are indicated in bold and the primers are included in boxes.

```
AAB88655 A. flavus      ------------------------------MKSDKDILMKPLPKSPGT 18
AAD43626 E. nidulans    MSPLETNPLSPETAMREPAETSTTEEQASTPHAADEKKILSDLSAPSSTT 50
AAB88660 A. fumigatus   --------------------------------------------------

AAB88655 A. flavus      G----------STTTGHSVSH-------------AEEVLDRQLHTPVSQI 45
AAD43626 E. nidulans    ATPADKEHRPKSSSSNNAVSVNEVDALIAHLPEDERQVLKTQLEEIKVNI 100
AAB88660 A. fumigatus   --------------------------------------------------

AAB88655 A. flavus      GFFGIYRYATRWDVAILFGSALAAIAGGAALPLFTVLFGRLTSTFQDIAT 95
AAD43626 E. nidulans    SFFGLWRYATKMDILIMVISTICAIAAGAALPLFTILFGSLASTFQRIML 150
AAB88660 A. fumigatus   --------------------------------------------------

AAB88655 A. flavus      HRITYDHFHHELTKNVVYFIYLGAAEFVAIYLATVGFIYTGDHVVQQIRV 145
AAD43626 E. nidulans    YQISYDEFYDELTKNVLYFVYLGIGEFVTVYVSTVGFIYTGEHATQKIRE 200
AAB88660 A. fumigatus   --------------------------------------------------

AAB88655 A. flavus      EYFQAILRQNIAFFDTLGAGEITTRITADTNLIQDGISEKVGLALTGLST 195
AAD43626 E. nidulans    YYLESILRQNIGYFDKLGAGEVTTRITADTNLIQDGISEKVGLTLTALAT 250
AAB88660 A. fumigatus   --------------------------------------------------

AAB88655 A. flavus      FVTAFIIAYIKNWKLALICSASLLALLLTMGGCSTLMLIFSKKALEYQGR 245
AAD43626 E. nidulans    FVTAFIIAYVKYWKLALICSSTIVALVLTMGGGSQFIIKYSKKSLDSYGA 300
AAB88660 A. fumigatus   --------------------------------------------------

AAB88655 A. flavus      GASMAEDILDSIRTVAAFNAQETLARKYESHLKDAEGPGMKSKVIFAIMV 295
AAD43626 E. nidulans    GGTVAEEVISSIRNATAFGTQDKLAKQYEVHLDEAEKWGTKNQIVMGFMI 350
AAB88660 A. fumigatus   --------------------------------------------------

AAB88655 A. flavus      GALLCIMYLNYGLGFWMGSRFLVEGISNIKAGDVLTIMMAIILGSYNLGN 345
AAD43626 E. nidulans    GAMFGLMYSNYGLGFWMGSRFLVDG--AVDVGDILTVLMAILIGSFSLGN 398
AAB88660 A. fumigatus   ------MRGIRSLPCWAPG----------------LSTKRIPPRELFAD 27
                              *   .*  *  .                 :   :     :.:

AAB88655 A. flavus      VAPNGQALSDAVAAASKLYGTIDRQSPLDALSDQGKTLEFVRGNIVLQNI 395
AAD43626 E. nidulans    VSPNAQAFTNAVAAAAKIFGTIDRQSPLDPYSNEGKTLDHFEGHIELRNV 448
AAB88660 A. fumigatus   LFPNACVISARHSARNGLIRQFSGCS---------------------GSI 56
                        : **. .::   :*   :   :.  *                    .:
                                                  Walker A
AAB88655 A. flavus      RHVYPSRPEVTVAHDLSCYIPAGKTTAFVGPSGSGKSTIISLLERFYDPV 445
AAD43626 E. nidulans    KHIYPSRPEVTVMEDVSLSMPAGKTTALVGPSGSGKSTVVGLVERFYMPV 498
AAB88660 A. fumigatus   SNSCNPRPYRSAITSLLSANVCSKGVSAVQP-------------RFLSTV 93
                         :   .**  :.  .:      ..* .: * *              **  .*

AAB88655 A. flavus      AGTIMLDGHDIQTLNLRWLRQQMSLVSQEPRLFATTIAENIRYGIIGSRF 495
AAD43626 E. nidulans    RGTVLLDGHDIKDLNLRWLRQQISLVSQEPVLFGTTIYKNIRHGLIGTKY 548
AAB88660 A. fumigatus   R---------LFSTSQRSLEPKSNVKSTGGQVVRPELHQDQEHEDIEKGF 134
                                  :   . * *. : .: *    :. . : :: .:  * . :
                                                                      ABC
AAB88655 A. flavus      EKESTYEIRKRVEAAARMANAHDFIMALPNGYDTNIES--FSLSGGQKQR 543
AAD43626 E. nidulans    ENESEDKVRELIENAAKMANAHDFITALPEGYETNVGQRGFLLSGGQKQR 598
AAB88660 A. fumigatus   ELSERAAQAAQVNLSAKLAKDG---------------------------- 156
                        * ..       :: :*::*:
                                   Walker B
AAB88655 A. flavus      IAIARAIIKDPKILLLDEATSALDTKSEKLVQAALDKASKGRTTIVIAHR 593
AAD43626 E. nidulans    IAIARAVVSDPKILLLDEATSALDTKSEGVVQAALERAAEGRTTIVIAHR 648
AAB88660 A. fumigatus   ------------------------------------------------AAGK 160
                                                                        * :

AAB88655 A. flavus      LSTIQKAYNIIVLANGQIVEQGPHEHLMDRRGIYCDMVEAHEIKKRYSRY 643
AAD43626 E. nidulans    LSTIKTAHNIVVLVNGKIAEQGTHDELVDRGGAYRKLVEAQRINE----- 693
AAB88660 A. fumigatus   KAGFKEIWRLLLIAR---------------------------------- 175
                         : ::   .::::..

AAB88655 A. flavus      SKRYSQLLTNLSPKHNPMTFFFDKDYPGDDESDIYSILSDDASDIGLHTG 693
AAD43626 E. nidulans    -QKEADALEDADAEDLTNADIAKIKTASSASSDL------DGKPTTIDRT 736
AAB88660 A. fumigatus   --------------------------------------------------

AAB88655 A. flavus      EKQRPVSRMSLSHLMQPVKEEAYSFWTLFKFLASFNRPEWPFLLLGLCAS 743
AAD43626 E. nidulans    GTHKSVSSAILSK-RPPETTPKYSLWTLLKFVASFNRPEIPYMLIGLVFS 785
AAB88660 A. fumigatus   -----------------------------------PEAKKLALAFLFL 188
                                                          **   : :.:

AAB88655 A. flavus      ILAGGIQPSQAVLFAKAVSTLSLPPLEYPKLRHDANFWCLMFLMIGIVSL 793
AAD43626 E. nidulans    VLAGGGQPTQAVLYAKAISTLSLPESQYSKLRHDADFWSLMFFVVGIIQF 835
AAB88660 A. fumigatus   LVSSGITMSIPFSIGKIMDTSTKATTEGGNELFGLSLPMFYGALAGILTL 238
                        :::.*   : ..  .* :.* : .  :  :  :  .. .:  :   : **: :
```

AAB88655 *A. flavus,* ABC multidrug resistance protein; AAD43626 *E. nidulans*, multidrug transporter; AAB88660 *A. fumigatus*, ABC multidrug resistance protein

A4.3 MFS Alignment 2

Clustal W alignment of six functional fungal MFS transporter genes. The motifs (D, C and H) are indicated in bold and the primers are included in boxes.

AAM48914 G. zeae, trichothecene efflux pump; *EAA72507 G. zeae,* mfs-multidrug-resistance transporter; *AAD127567 F. sporotrichioides,* trichothecene efflux pump; *AAS66020 A. parasiticus,* potential toxin transporter; *CAF31980 A. fumigatus,* AflT-like major facilitator superfamily protein; *AAW03302 A. fumigatus,* multidrug resistance protein 3

Appendix V Complete Sequence of Phagemid Insert with gABC 4e

6725 bp of sequence from the *N. lolii* Lp19 genomic DNA Lamda library. The Open reading frames, gABC 4e (3866-5797) and the Orotate phosphoribosyl transferase (995-2077), are underlined. With the sequencing primers from Table 2.4 indicated in bold.

```
   1 TTTAAAAAGC GAGCCTCCCT GTTCAGAGAG CGCACCAATG ATGAGACTGG
  51 AGTAGTGGGA TCGAGAGCAG TAGGGTGGAT ATCGTGGCGA GGTTTATTAT
 101 CGAATTATTA GTTCGCCACA TGCTCCGTAC GGGGCGCCAA CCTCGACGAG
 151 CGTCAAGTCA GGGTGCGACC CGACAGGCTG GACCGTTTGG TGGACATTTT
 201 GCTTTCGGCT CAGTGTCAAT GAAGATTCAT ATGCGGTATT TCTGCACCCG
 251 GTGTTAACGC CTCGGCCACG CAGTCGTTTG GATGCAAAGG AAGCCTTCCG
 301 CGCGTGCGGG TCGGCCGCCA AACCCTCAGG AAAATATGCC GTGTAAAGCT
 351 TCACAAAGTA CAGAATACGG AATACAAGCA TCAGAACCCG ATGTTTGTAG
 401 GTGAGCTTGA CTTCCAMACA GCCAGAATCG AATTTATGTA AGATTGGCTG
 451 CAAGGCACAA GTTGGCACGG GCCTTCAACC AAGGTAGTGA GTGCGGCTTC
                ABCend_F3a/b
 501 TGTTCTKGGT CCCTGTTGCA GTTGARTTGC AATTTTCATK GCGGCATCGG
 551 TTGACCCTGT TTYGAGGCCT TCTGATGGCT GCTCGAAACG AGCTGARACW
 601 YGYTCTTCTA CCGAGTAAGT TCTTTCTTTT CATGGGTGAT GTGGTCCAGG
 651 CCTTTTGGTG ACGCCCGTTG ATTCTGCAAC AAGTCCGGAG TACGGAGTAG
 701 AGCAAGTCGC CTCGTACTAT CAAATGCTGG TTGCCAATCT CCCAAGTCGA
 751 GCAAGGACCT GATTGCAATT TTCCTATCAG TTGCAGAAGG AACATGTTTT
 801 CTTTCGTCTT GCATTTCTAC CAATAAGGGG CGAAGGGTTC GAAGGAACAG
 851 AATGCACGGC TCGATTCGAA TTCAAGCCAG CAACACAACG TGACACAAAC
 901 TCGGGCTGGT ACAACATGGT GGGTTTTCAG CTAGCCGTAG CTATTGAAAG
 951 AGCGAGAATC TCACACAGTC TAATATTCCA GATAAACAAG CTTCCTAGTC
1001 CTTAGGTTCT ATTCTGCCTA AGCTAGACTT AGGAAAGCAA AAGAGCCTAG
1051 AGTATTTACT TCTAAAGCGA AAAGCCTATT ACTAATTACT ACTAACTTAC
1101 TATGCATAAT ATAGACTAAC AAAATATCTA ACTAAGCGAG RGAATAAACG
1151 TTATAACTAA CTAAACGTGC CTTAGCTATA AATCGCACTT TTTACAGTCT
1201 TAGCTAAMCA AAGCACTTAT ACGAAGTCAG ATGTTTAACG AAGTAGATCT
1251 TATGCTGGTT TTGTTTTTAG TACATGTTAG AAATTTTGTT ATCCCAAAGG
1301 TAAGTTCCAC TATTAGTCTA CGTCMATAGC CGCGCCTAAT TCCCCGCCGT
1351 TGATAGCTGA AAACTTCCTG TCTAACTAGT CCGTTGCCTT TGTACTTCAG
1401 ACRATAMTAA WWCAGTTCGT TTGATATCGT CCTCAGAGGC GAAGCTCTTC
1451 ATGCCCTCAA TGATGTCATC CAGAGTCAGA ATTGCAAAGA TTGGGATGCC
1501 ATACTCTTTT CTCAGTTCGC CAATGGCGCT AGGACCAGGT TTGGAGTCGT
1551 CACTATCGGC AGCAGGCAGC TTCTCCTTGC GATCCAGGGC AACAACAATA
```

1601 CCTGAAACAA TGCCGCCCTC CTTTCTAATC TTATCGATGG CTTCTCTCTT

1651 TGCAGTTCCA GCTGTGATAA CATCATCGAC GATCAAAACT TTCTTCCCTT

1701 TCAAAGAGGA ACCAACAATG TTGCCTCCTT CACCGTGGTC TTTGGCTTCT

1751 TTCTGTCARR CGAGTAAGAG ACSGTATCCA GATYTTGAGG GCAAGCTCGC

ABC4e_F12

1801 CGAGCCTGAT GGTGATGGCG GAGCACAGAG GAATACCCTT ATACGCAGGA

1851 CCAAAGACAA TATCGAAATC GAGGCCATCA TTCTGCTGGG CATCAAGGAT

1901 GGTCTTCGCG AATGCGGACG AGATGGCACC GGCGAGACGA GCGGTATGGA

1951 ACTCGCCCGC ATTAAAGAAA TAAGGSGAGA GTCTCTTGGA TTTGAGCTCA

2001 AAGCTACCAA ACTTCAAGAC TCCGCCCTCA ATGGCGGCCT GGAGGAATTC

2051 CTTTTTGTAG GGAGCGAGGC CGGTCATTTT TTTCTTTCTT TCTTTCTGTT

2101 TTGCCTGGTT TTGGTTCAGA ATGATTGGTC CTGATGGATG CGATTGTGGA

2151 GGAGGTAGTC CAGCAGTAGA AAATCACCCC TTTTGTTGTA ACAAGCCTTC

2201 CGCCGTCGGT CTCTTCCAGG CAACCTGGCT TGTAAATCAA ATATAGTGAA

2251 GTCGTTGCAA TGTGGACTAC AGAGGAAATG AACCAAAAAA GCTTTGCATC

2301 CAAGATGTTG TTCAATGGTT CAATGGGGCG CTAAATCGGC GAGGAGAGGA

2351 AAAAAAAAGG CACAAGGGGG GGAGCCTTCG CAATGACTTA CCGAAATTGA

2401 TATTTTTCAA GTGGCAACAT CAGCAGACCA AGCGGTCCGA TGCAAAATTG

4eABC_F10

2451 TCGTTCTTCA TGCAGCGGCA GCGATGACAT GCGCTGATGT TGTCGAAAGC

2501 TCTGTCAGTA GTCCTTTGTC GTCTTCAAAT CAGTAAGCTG TTTCTTCCTT

4eABC_F11

2551 GTGCAGAGGG AGTGGGCGAG TGGGTGGGCA GGCTGGTGCA ATGCTCCTGC

2601 AAGAATGCTG AACGCACAGG AAATGAGGCT TTTGATGAAA TTTGAATTGA

2651 ATTCTGACTG TGGTTGTGAT GATGGATCAT CTGCAAAAGC TCTTGATGAA

2701 AAAAATTATT AGGTGGGGAT TTTGTGTTAG GTGGGGACAG GTGCCTCGCA

2751 CCCCCCCCCT TTTCACCTGC CAGGGTCGCC TTTGAATGCT GATTAGTCAC

2801 TCGGTCGCAC ACCCAATAAT ATCTTGCGGC AAGTTAATAT TACCTTGGAG

2851 GTCGAAAATG CTTTCCCCGA CGCCAAGCTC GGAATCTTCA GTTCCGACAA

2901 TCCCATTGAG CCAAATGTCT CATCACTCAT CACGCCTTTT AGTGCTAAGC

2951 TGCAAGGAGT ATTTCGCTGA CAAGACAACC TCCGAACGCA CGAGTCTGAC

3001 CCAGAGGGCG CTGTGGCTTG AGGTTTTTGC CCTGCCTTGC TTGCCTTAAC

4eABC_F7a/b

3051 CTAGGGCCAT AGAGTCTCGA GACGGGAAGA CGAGCACGTT TTCAAAGGAG

3101 TAATCCGTAC CTGCAAGATC AACMAGATCC AGTCAAGGGT ACCCGGAACT

3151 AAGGGATGAC CGGCTGTCAT CTTGACCGCT CCGACGCGAC AGACCGTTCA

3201 GCCGCTGTAG CACGTACTTG CGGAACAATG ATGAACATGA CGACTTGGCA

3251 GCATCATTCT GCCTAGACAG GCAGGTGCGA ATGCTTAGAT TATCGACTTA

3301 TCCCATCCCA ATGATTGTGA CCACTCGGAA ACACTGTTTT CAGGCTAAAC

3351 TTTAGGTAAC AAAATGAACA GACTCAACGT CGTGGGCGAC GGTCGACGGG

4eABC_F6a/b

3401 CAGGAATAAT ATGTATACCC GCATCCGCGG TAAACATTAG CTGCAATTGG

3451 ATACGTGCCT TCCGCAAATG CAACAAACAT GCATTTATCA GGGGCACCAG

3501 TAATAAATGA CATCAACAGG CGCATTGCTA CCGACAGTTG GACATGCAGC

3551 TGCGCGAATA AAATACATCG GCATCATTCT ATATGTACTT TCGTCGTCCT

3601 CGTCAAATAG TCGGAGCAAG CATAGAATTT CGGTACACAT TTGATGTGAC

3651 CCTCGCGACT GATGCATGGA CATCCCCTGC AATTGCAGCT CAAGGGAAAG

3701 ACGAAGCCCG CGGCTTCCAG TCCATTCAGA TTGCAATTTG AAGACGGCAT

4eABC_F5a/b

3751 GTGCCAGTAC TGGTCCATCT ATATCTATGC TCGGAGTATG CAGGGCTAGG

3801 TGTTGTGCAC TTGAGTCGGC GTTTTGGGCA CGAGTGCGCA TCGCGAAAGG

3851 AAGAAAGCGC CAAAAATGTC TGGCCTCTTG TTACACCTAG ATGCTGTAGG

3901 GAGCCTGATC TGTCACAACC CGAGAGAAGA TTTGATATGT GATAGCCAAG

3951 TGCATTTGAC CCGGCCACCC AGAGTTTGTT TGACCTCGAT CCAGCATTCC

4eABC_F4a/b

4001 GTCGCTTTAG GCCGCGGTGA GGGAGGGTTC AGGTCCATCG GAAGCTTCCA

4051 TATTTCGGTC GACAACGATC CCCTAAAGCT AACACACCGG TTCTAGACAA

4101 CGTGCGGCTG GTCACATTGG GCTTCTATAG CTGGGATTGA TGAGACTGCA

4151 AGCAGCGATG GACTTACCAG TCTGACATTT ACGTACGCCT TGTTTGGACT

4201 CGGCCGCAGC ACTCTGCAAC CCAGCCCAGT CGCAGCCGCC GGCCTTTGCC

4eABC_F8a/b

4251 ACGCATCACG GTAAGTCAGT AAGTAGCCGC CGCGTTATCT GGCTTCTTAT

4301 CAAATTCTAT AAGCGACTTA ACTGGTCTGA GTGGCAGCGC CTTGGGCGTT

4351 ATTCTTATCT GCATTTCCAC GCTTATATCA GGAGTAGTCG TTGCGCTCGT

4401 CCTTGGTTGG AAACTGGCAC TCGTGTGCTT TGCCGTTATT CCTTTGATGA

4451 TCGGCGGAGG TTATTTTGGA GTTTCTCTTG TTAGCAACTT CGAGAAGAAG

4eABC_F3/b

4501 AACGAACTAT TTGCTAATAA GGCCTCGGAG TTTGCTGGCG AGACGCTGAA

4551 TGGTATACAA ACTATCGCGG CGCTAACCAA GGAGCGGACC GCATTAGCCG

4601 AATTCGAGGA AATACTCGGC GAGACTAAGA AGGAAGCTTT ACTCGCAAAC

4651 TTGCAAGCTT CCTTCATGTA TGCACTAACC CAGTCAGCCT ACTATGCCTG

4701 TATGGCGCTG AGCTTTTGGT ATAGCGGGCG GCTGATCCTT GACGGCGAGT

4eABC_F9a/b

4751 ACACTTTGTT CCAAGCCATC GCTATCCAGT CAACCATGCT CCTGAGTGCC

4801 TACTCGGCTG GCGTTGTCTT CTCCTGGACT CCTAACATCG GAAAGGCTAA

4851 ACAGGCTGCT GCCTCACTAC AACGACTGCT AGCATTGAAA TCTGCAATCG

4901 ACCCCTCGTC TCCGTCTGGA GAGGGAATTA CCTCTATGCG CGGCAATATC

4951 ACGTTCGACT CGGTTAGCTA TTCATACCCT TCCCGCCCGG ACTACCTGGC

4eABC_f2

5001 GTTGGACAAT GTCAGCTTCA ACATTCCAGC AGGCGCCAAT GTCGCCTTTG

5051 TTGGCACCAC AGGTTCTGGC AAGAGCACCA TCGTCTCGCT GATTGAGCGG

5101 TTTTATGACC CAACCAGTGG GAGGATACTT GTGGACTCCA AACCTGTTAC

5151 GTCGCTCCGC TTGTCCGAAT ATCGAAAATG CATTGGTTTG GTCAGCCAAG

5201 AGCCCACCCT GTTCAACGGC ACCGTCAAGA TGAATTTGAC GATTGGTTTG

4e_F2b

5251 GACGAAGAGG GCCAAACACC ACCCACTAAT GAGGAAATCG AGAATGCGTG

4eABC_r2/b

5301 TCGCTCAGCC AACATTCATG AATTCATTAC CAGCCTTCCG TGAGTACTCA

5351 AATTTCCATT ATCTTGTCTT TTGGTCACAA ACAAAAACCA TATCATCATA

5401 CCCTCTGCCT ACAAATCGCT GACGAACTCT GTTCGATAGG GATGGCTACG

5451 ACACGGAAGT TGGCAGTAGA GGCAGTCAGC TCAGTGTCGG ACAAAAGCAG

4eABC_R4a/b

5501 CGTGTGGTGC TCGCCCGGGC CCTGCTTCGG CGACCTAAGG TTCTGTTGCT

5551 CGACGAGGCG ACGTCTGCGC TGGATTCGCA ATCAGAAGCC AGCATTCAAC

5601 AGGCGCTTGA GCTGGCTAAA GAGGGTAGGA CAACAATCAC GATTGCGCAT

5651 CGACTGAGTA CAATTGTAAA AGCGGACAAA ATTTACGTCA TGGGTGACGG

5701 CAAAATTGTC GAAGCGGGAA CACACGCTCA GTTGATGGCC AAGAAGGGTG

4eABC_R3/b

5751 CTTATCATGG ACTGTACGTG GCAAACAAAA GTGGGCAAAC TCTGTAAATG

5801 TTGTGAATGC TGTAGATGTG GCAGATTATT ATTGTACCAA TTCCAACATA

5851 TCCTGAGACT GTGGTCAGGT TATGTCAGTC CAATATAATC GAGGTACAAA

5901 CGCAGAACAT TGTCCTCGCC CTGTAAGCGA GGCTGTTGGC ACGGTTTAGG

5951 CTCTTTTTTT TTTCTGTCTG ATAAGGCGAC AATGACCAAT TCTGGTGAAT

6001 GGCATTGATG GCAGCTCAAT AAATTTACGT ACGGATTACA GTTGTTGCTG

4eABC_R5a/b

6051 GATAATAAGT GTTGATAAAC ACCACAGCAT ACTCATAAGT GGTGACATGA

6101 GCAAGTCAAA TCGAAACCGC CTTTAAATTT GTTGGTATGC AATGAGGGTT

6151 TCAGAAGGAG TGCAGCCAAC AACCCGGCAT CTACAACTCT CCATGGTGCG

6201 TTTGCCCCAC GGAGCCAGCT ACAGGGTTAT TAGCAGGAGA CTGGAATGAA

6251 TGTGGCCCCG CGGAAATGCT GACGGCTGAC GGAAATGCTG ACACACACCT

6301 GGTGGATGCA CCCGTCCGGA GGACTCAGCT TTGTGTAAAA AAAAAAAAAA

4eABC_R6a/b

6351 AAAGCTGACT CCTGCTGACA CACACAGATG AGCTTGATAA TAAGTGCCAG

6401 TTACAGGGCC CCCTGACGG GCCTGACGGG CCTGACGGGG ATCAACTGAG

6451 CCAATTCCTG ATATGGCAGC TCGTTGGCTT TTGTTCCGGT GGTGGCGCCG

6501 CCCCCCCCCC CCGGCACCAG CAAATCAAGT TCATTGTGGG CTGGTGTGAA

6551 TGGGAGTGAG AGTGACATCG GGGTGACTTC TTGTTGGGTG ACAATTCAGC

6601 CCGGTATCAG TGTAGATACG AGTCTAAACC GGTCCGACAC AACGGCCCAT

6651 CCTGCGAATG TATCCCCTGC GTCCCGGTCC TGACCTCCCA ACGCTCGCTG

6701 CCCACTCCCG CAGGAACGTA GAATA

Errata (September 2006)

Page 6. Section 1.2.2.1 Lolines
N. uncinatum infects medow fescue not tall fescue

Page 106. Figure 3.27.
The position of intron 1 as shown is not thought to be correct and requires further investigation.

Page 25. Table 2.1 Strains and Plamids
Fungal Endophyte strains

Fungal Endophyte Strains

Epichloë festucae		
FL1	Host: *Festuca longifolia*	(Leuchtmann *et al.* 1994)
Epichloë typhina		
E8	Host: *Lolium perenne*	(Schardl *et al.* 1994)
Neotyphodium coenophialum		
RS2	Host: *Festuca arundinacea*	(Christensen *et al.* 1998)
Neotyphodium lolii		
AR1	Host: *Lolium perenne*	M. Christensen (Agresearch, NZ)
AR40	Host: *Lolium perenne*	(Christensen *et al.*, 1993)
AR48	Host: *Lolium perenne*	(Christensen *et al.* 1993)
AR66	Host: *Lolium perenne*	(Christensen *et al.* 1993)
Lp19	Host: *Lolium perenne*	(Christensen *et al.* 1993)
Neotyphodium uncinatum		
Fp2	Host: *Festuca pratensis*	(Christensen *et al.* 1993)
FATG-2		
TF15	Host: *Festuca arundinacea*	(Christensen *et al.* 1993)
LPTG-2		
Lp2	Host: *Lolium perenne*	(Christensen *et al.* 1993)
LpTG-3		
AR501	Host: *Festuca arundinacea*	(Christensen *et al.* 1993)