The indole-diterpene gene cluster from the ryegrass endophyte, *Neotyphodium lolii*, is required for the biosynthesis of lolitrem B, a bioprotective alkaloid.

This thesis is presented as a partial fulfillment of the requirements
for the degree of
Doctor of Philosophy (PhD),
in
Molecular Biology
at
Massey University, Palmerston North,
New Zealand

Carolyn Anne Young

2005

Abstract

Lolitrems are indole-diterpene alkaloids produced by *Epichloë* and *Neotyphodium* endophytes in association with their host grass *Lolium perenne*. Some indole-diterpene (ID) alkaloids are proposed to have insecticidal properties, but lolitrem B is known as the causative agent of the animal syndrome ryegrass staggers. Lolitrems are preferentially synthesised *in planta,* which suggests that the genes required for lolitrem biosynthesis are symbiotically expressed.

The lolitrem biosynthesis pathway has been proposed as a metabolic grid based on the identification of likely intermediates from endophyte-infected ryegrass. Closely related ID compounds are expected to serve as substrates for the same enzyme, but until recently these steps had not been validated. The identification and characterisation of a *Penicillium paxilli* gene cluster required for the synthesis of the ID paxilline has identified key enzymes required for the production of the ID backbone. Based on the similarity of lolitrem B to paxilline it was proposed that these two biosynthesis pathways would share orthologous early steps but later steps to convert paxilline to the more complex lolitrem B would require additional enzymes.

The lolitrem biosynthesis genes (*ltm*) were isolated using degenerate PCR and from candidate genes identified as ESTs in cDNA libraries. Ten *ltm* genes were identified that had functions consistent with those required for lolitrem B biosynthesis. The 10 *ltm* genes were contained on three gene clusters that are separated by repetitive AT-rich sequences that contain remnants of retrotransposons. The *ltm* clusters 1 and 2 contain eight genes, seven of which are orthologues of the characterised *P. paxilli* paxilline biosynthesis gene cluster (*pax*). Functional characterisation of *ltmM* an FAD-dependent monooxygenase and *ltmC* a prenyl transferase confirmed these two genes were required for ID biosynthesis and were orthologues of *paxM* and *paxC*, respectively. All 10 *ltm* genes have similar expression profiles and were highly expressed *in planta* where the production of lolitrem B is most prevalent. The taxonomic distribution of the *ltm* genes has established which endophyte strains are likely to produce ID compounds. This work provides the basis for elucidation of the lolitrem biochemical pathway and opens the way for determining how the plant regulates the synthesis of this important group of bioprotective molecules.

ACKNOWLEDGEMENTS

Embarking on a PhD was not an easy decision for me. I had family, I had never studied full-time, I was older than the average student and was it really something I needed to do? I finally made the decision to go for it, handed over my managerial role and I became a student!

Life has an interesting way of changing your focus by throwing a curve ball every once and a while. Mine came as a little 7lb 13oz package at about the PhD midway stage. He wasn't quite the experiment I had planned but we named him Oliver and love him dearly. This PhD could not have been completed with out the support of four very important people. The first three are my family, David, Patrick and Oliver who had the responsibility of making sure I had all the time I needed to do what I needed, even when I said I didn't need that much time. They fed, me they looked after me, they picked me up when I got low. They were there whenever I needed them. The fourth person is my supervisor Barry Scott, who feels like family. He supported my PhD decision, my pregnancy, my research and so much more. Just saying thank you to these four people is not enough. I am truly grateful for their love, their support and their encouragement.

Andrea Bryant was just the person I needed in the lab. Andrea made me smile, she looked after my plants, she kept me sane, she helped me when I needed an extra pair of hands and was unflappable when we were working together.

Thank you Geoff Jameson for correcting me. After you did such a great job on my Masters thesis I was delighted when you were prepared to read my PhD thesis. I hope that my writing style and grammar has matured for the better.

Emily Parker was invaluable for discussions on indole-diterpene biochemistry. I knew how I wanted to place the place the catalytic functions and Emily made sure they were realistic.

Thank you to Russell Poulter and Margy Bulter who gave me great advice on retrotransposons, especially Russell whose seminar on retrotransposons made my AT-rich regions make sense.

I would like to acknowledge Pete for his evolution expertise and the trees that have been used in Figure 3.60.

I would like to acknowledge the support I received from my co-supervisor Greg Bryan (AgResearch). I would like to thank Mike Christensen, Wayne Simpson and Anouck De Bonth for their help and guidance with all things plant like. You were my green fingers and I really appreciated it. I would also like to thank Brian Tapper and Liz Davies for the HPLC analysis and their input into what the peaks meant.

I would like to thank Richard Johnson (AgResearch) and German Spangengberg (Agriculture Victoria) for supplying me with some of their interesting EST sequences. I believe that the data generated with their sequences was the icing on my thesis.

Thank you to all past and present members of the Scott Base lab. Lisa, Andrea, Austen, Rohan, Brendon, Michelle, Kim, Aiko, Simon, Xuiwen, Shuguang, Christina, Emily,

Richard, Mike, Grant, and others that I may have missed. Some of you had been and gone before I started my PhD but you have still influenced my scientific career.

Thank you, to Rosie, Max, Barbara, Pat (all of them), Liz, Paul, Isabel, Chris, Catherine N, Catherine D, Kathryn, John, Gill, Trish, Lorraine, Pete, Cynthia, Neville, Vikki, and many other members of IMBS (both past and present) who have given me friendship, encouragement and support.

Lots of little thank yous are needed for; Joe Win for his advice and a printer so my thesis could be printed on A4 paper while living in the US; Brendon for the chemical structures; Chris and Mike (and many others) who entertained David while I was working; Kathryn Stowell and Robyn Marsden for their patience with the developer, making sure it was working when I needed it most; Mrs C for Oliver transport; Sanjay for the *P. paxilli* HPLC analysis; John McKay and Kirsty Allen for help with the lightcycler.

I would like to thank my Ballroom Dancing group, Leigh, Lynne and Pete, Smurf, Shirley, Wendy, Kevin, Jacinta and Bill, Vicky, Yvonne, David and others. I enjoyed the dancing, I enjoyed the coffee, but most of all I enjoyed making new friends. A special thank you to Smurf and Leigh who were required to dance with me week after week.

Thank you to my dearest friends Kate, Kirsty and Sharon for their friendship, love and support. Also thanks to my friends Rich and Linda, Mike and Kate, Kirsty and Marcus, Mark and Vanessa, Sharon and Tony, Trish and Pete, Geoff, Marion, Austen, Emily, Lisa, Rohan, Rachel and others that I may have forgotten to mention.

To Louise and all the Outrims, Olly, Scott, Kenton and Liam. You were there for Patrick when ever I needed you. Then you were there when we REALLY needed you. Thank you for always being there.

After the arrival of Oliver, my PhD could not have advanced without the wonderful childcare from the Hoiho Centre at the Massey Childcare Centre. Oliver was so little but I knew he was in good hands. You mothered him, you babied him, you loved, you did it just right. I am especially grateful to Fleur, Anne, Raewynne, Kelly, Libby, Heather, Megan and the others who were there for Oliver when ever he needed you. Oliver would like to thank you as well and include his friends, Diana, Max, Brodie, Trent, Neitana, Jenny and many others.

I am grateful for the funding I received throughout my PhD. I was initially funded by a FRST Bright Futures Top Doctoral scholarship and then in my final year by the Bio-Protection CoRE (Lincoln). I would also like to thank IMBS for financial support for conference attendance.

Finally, I am extremely grateful to David 'the twist and turn' who was my rock during the printing process.

It is hard to believe that the end is near and it is time to move on for new life experiences. I shall miss my project but look forward to talking to those who contribute to this work in the years to come.

TABLE OF CONTENTS

List of Tables

LIST OF FIGURES

Abbreviations

bp	base pairs
kb	kilo bases
kDa	kilo daltons
pfu	plaque forming units
Mb	mega bases
ACM	Aspergillus complete media
ACP	Acyl carrier protein
AFLP	Amplified fragment length polymorphism
AT	Acyltransferase
BAC	Bacterial artificial chromosomes
BLAST	Basic Local Alignment Search Tool
cDNA	copy DNA
CD	Czapek Dox
CDYE	Czapek Dox + yeast extract
DH	Dehydratase
DMAPP	Dimethylallyl diphosphate
DMAT	Dimethylallyl tryptophan
DNA	Deoxyribonucleic acid
ELISA	Enzyme-Linked Immunosorbent Assay
ER	Enoyl reductase
EST	Expressed sequence tag
ETC	Epichloë typhina complex
FAS	Fatty acid synthesis
GGPP	geranylgeranyl diphosphate
HCl	Hydrochloric acid
HMG CoA	Hydroxymethylglutaryl coenzyme A
ID	Indole diterepene
IPCR	Inverse polymerase chain reaction
IPP	Isopentyl diphosphate
IPTG	Isopropyl bD-thiogalactopyranoside
KR	Ketoreductase
KS	Ketoacyl synthase
LB	Luria broth
LINE	Long interspersed nuclear elements
LAE	Lolium-associated clade
LTR	Long terminal repeat
MFS	Major facilitator superfamily
NRPS	Non-ribosomal peptide synthase
ORF	Open reading frame
PCR	Polymerase chain reaction
PD	Potato dextrose
PKS	Polyketide synthase
RAPD	Random amplified polymorphic DNA
REMI	Restriction enzyme mediated integration
RIP	Repeat induced point mutations
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RT	Reverse transcription

RT-PCR	Reverse transcriptase-polymerase chain reaction
SINE	Short interspersed nuclear elements
TAGKO	Transposon-arrayed gene knockouts
TE	Thioesterase
Tris	Tris (hydroxymethyl) methylamine
TLC	Thin layer chromatography
Tris-HCl	Tris (hydroxymethyl) methylamine pH changed with HCl
X-gal	5-bromo-4-chloro-3-indolyl-bD- galactoside

CHAPTER ONE

INTRODUCTION

1.1 Fungal endophytes of ryegrass

Epichloë endophytes are an important group of clavicipitaceous fungi (*Clavicipitaceae*, Ascomycota) that colonise the intercellular spaces of temperate grasses from the subfamily Pooideae. Endophytes act as biological agents for growth and persistence of perennial ryegrass pastures by conferring protection from drought, insect pests, and fungal and nematode pathogens. The endophyte gains nutrients from the host apoplast and the endophyte/ryegrass association facilitates dissemination of the fungus. It is within this endophyte/ryegrass association that the endophyte is able to produce a range of secondary metabolites with known anti-fungal, anti-insect and anti-herbivore properties. Consequently, these bioprotective secondary metabolites are proposed to enhance not only the fitness of the fungus, but also that of the grass (Clay, 1990; Clay, Schardl, 2002).

During the vegetative phase of growth, the sexual *Epichloë* species form asymptomatic associations. They predominantly grow between the intercellular spaces of the ryegrass with limited epiphytic growth. Floral development triggers a switch of the fungus to epiphytic growth and the formation of a stroma that chokes the emergence of the floral meristem, thereby causing infertility. The heterothallic mating system of the sexual *Epichloë* is mediated by *Botanophila* Séguy flies, which are attracted to the stroma and spread the spermatia to that of the opposite mating type (Bultman, et al., 1995). The mature stroma releases meiotic spores that germinate on a grass floret and infect the host ovule (Chung, Schardl, 1997). Upon germination of the seed the endophyte once again colonises the intercellular spaces and the cycle can be repeated. The highly antagonistic *E. typhina* species can only be acquired by another plant via horizontal transmission. Some *Epichloë* species, such as *E. festucae*, form pleiotropic associations where some tillers give rise to stroma that choke and some tillers remain asymptomatic. These species can be acquired by both horizontal and vertical transmission. The asexual *Neotyphodium* species form a totally mutualistic association with their host, where they remain asymptomatic in the plant. These asexual endophytes are able to grow into the developing seed and are then transferred vertically without sexual recombination (Philipson, Christey, 1986; Majewska-Sawka, Nakashima, 2004).

The *Epichloë* genus consists of 10 known distinct sexual species (known as telomorphs). However, fungal taxonomy classifies the asexual species (known as

anamorphs) in a separate genus, called *Neotyphodium* (Glenn, et al., 1996). Molecular phylogenetic analysis using sequence data from the genes *tub2* (Fig. 1.1), *act* and *tef* (Schardl, et al., 1994; Tsai, et al., 1994; Moon, et al., 2000; Moon, et al., 2004) and microsatellite analysis (Moon, et al., 1999) supports the hypothesis that the asexual *Neotyphodium* species are derived from *Epichloë* spp. The most closely related ancestors of *N. lolii* are the *E. festucae* (Fig. 1.1) (Moon, et al., 1999). Further analyses of the *Neotyphodium* species have confirmed that most are interspecific hybrids. They are asexual and have been found to have up to three different *Epichloë* parents as can be seen from the presence of multiple copies of genes whose sequences align with those of the related sexual species (Fig. 1.1) (Schardl, et al., 1994; Tsai, et al., 1994; Moon, et al., 2004; Collett, et al., 1995). Interestingly, there appears to be only one parental rDNA maintained in all asexual hybrids examined to date. The *Neotyphodium* sp. LpTG-2 strain Lp1, an interspecific hybrid that contains ancestral DNA from an *E. typhina* E8-like strain and *E. festucae* (Schardl, et al., 1994; Collett, et al., 1995), contains rDNA of a single type that is identical to that of *E. typhina* (Ganley, Scott, 1998; Ganley, Scott, 2002). Therefore, the *E. festucae* rDNA was either lost or homogenised and replaced with that of the *E. typhina*.

The average genome size of the haploid sexual species and the interspecific hybrids has been estimated by quantitative Southern analysis and electrophoretic karyotypes (Murray, et al., 1992; Kuldau, et al., 1999). The chromosome numbers range from 4 - 13 with chromosomal DNA sizes from 0.4 to at least 10.8 Mb. The haploid strains such as *E. typhina*, *E. festucae*, and *N. lolii* were estimated at ~28 - 35 Mb (Kuldau, et al., 1999). As expected, the interspecific hybrids have a larger genome of up to ~62 Mb, consistent with the presence of multiple copies of genes. Lp1, an interspecific hybrid of *E. festucae* and *E. typhina* ancestry is ~55.8 Mb, whereas *N. coenophialum*, with ancestors from the *E. festucae*, *E. typhina* complex (ETC) and *Lolium*-associated clade (LAE) is ~61.2 Mb (Kuldau, et al., 1999). The Lp1 genome size is almost the sum of the two ancestral genomes but the *N. coenophialum* genome, consisting of three ancestors, is smaller than expected presumably due to loss of some chromosomes. This is supported by sequence data where three alleles are not always detected for each locus within *N. coenophialum* (Tsai, et al., 1994; Moon, et al., 2004).

To date few regions of the *Epichloë* genome have been sequenced. However, a 43 kb region of an *E. festucae* genomic library was recently sequenced and showed that small

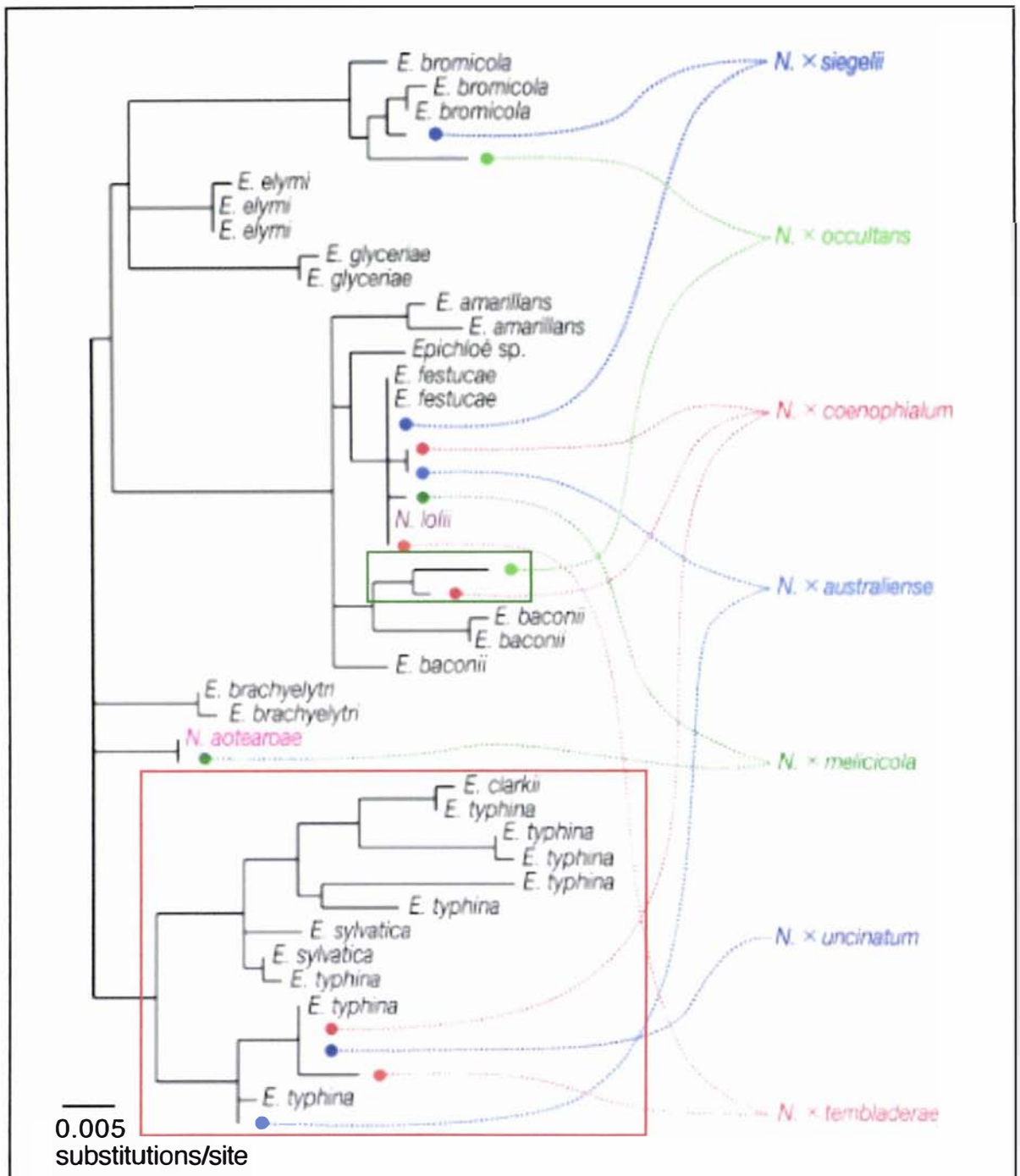


Figure 1.1 Phylogenetic Relationship of the *Epichloë* and *Neotyphodium* endophytes

The phylogenetic relationship of the *Epichloë* and *Neotyphodium* species based on maximum likelihood analysis of intron rich *tub2* sequences. The asexual species are indicated to the right of the tree with coloured lines linking to their ancestral parents based on the presence of multiple *tub2* alleles. The LAE-clade is boxed in green, and the ETC-clade is boxed in red. This figure was reproduced from Schardl and Craven, (2003) Molecular Ecology 12: 2861-2873 (with permission from Blackwell Publishing).

regions of microsynteny to the sequences of *Neurospora crassa* and *Magnaporthe grisea* were observed (Kutil, et al., 2004). Whole genome analysis of the *E. festucae* BAC library showed that only 20% of the clones contain repetitive DNA based on hybridisation of the library to total *E. festucae* genomic DNA (Kutil, et al., 2004). Sequence analysis of a clone represented only three times in the BAC library revealed regions with homology to reverse transcriptases, a domain associated with the *pol* gene from retrotransposons, and to other regions with greater than 70% AT content (Kutil, et al., 2004). Thus far there are few reports of transposable elements. There are, however, numerous microsatellite sequences within the *Epichloë* species (Moon, et al., 1999; van Zijll de Jong,E., et al., 2003). Three linear double-stranded DNA plasmids are found in some *E. typhina* (Mogen, et al., 1991; Chung, et al., 1996), but none of the three plasmids are located in the nuclear genome (Mogen, et al., 1991). At least one of these plasmids contains sequence similarity to a reverse transcriptase from *Neurospora* spp. suggesting the presence of retrotransposons.

Fungal biomass *in planta* has been determined by estimation of the hyphal volume present in a section of grass and by quantitative PCR. Both estimates from mature leaf sheaths resulted in fungal biomass to plant matter ratios of 1:200 - 1:500 (Panaccione, et al., 2001; Tan, et al., 2001).

The endophyte does not elicit any apparent host defense response and remains within the intercellular spaces of the host growing predominantly with in the leaf sheath and the aerial structures of the plant. Hyphae are not found within the root system of their plant host but originate from the meristem (Christensen, et al., 2002). Metabolic activity of the endophyte has been monitored in planta using a GUS reporter under the expression of a constitutive promoter (Herd, et al., 1997). The area of the grass with most fungal activity is the leaf sheath with endophyte metabolic activity decreasing as the plant increases in size.

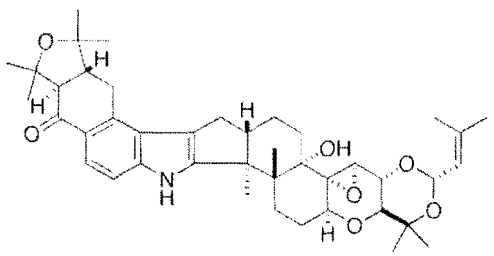
Analysis of proteins contained in the intercellular fluids from endophyte-infected grasses has resulted in the identification of several secreted fungal proteins, a subtilisin-like protease, an invertase and a ß-1, 6 glucanase (Lindstrom, et al., 1993; Lindstrom, Belanger, 1994; Reddy, et al., 1996; Moy, et al., 2002). The fungal proteinase, Atl, found in the association between *Poa ampla* (a perennial bunchgrass) and *N. typhinum* is present as 1-2% of the total leaf-sheath protein (Lindstrom, Belanger, 1994).

Orthologues of this gene have been found in a range of other endophyte species but their levels of expression are lower than that seen in the *Poa ampla*/*N. typhinum* association (Reddy, et al., 1996; McGill and Scott unpublished data). A ß-1,6-glucanase was up-regulated in the presence of ß-1,6-glucans and is also detected in the apoplast of endophyte-infected plants (Moy, et al., 2002). However, as ß-1,6-glucans are not present in plant cell walls the enzymatic activity of the ß-1,6-glucanase must be limited to components produced by the endophyte. The high level of expression of these genes during endophyte infection suggests that these hydrolytic enzymes are important in the interaction between the grass and endophyte. Hydrolytic enzymes such as the three mentioned above could have roles in nutrient acquisition, disrupting protein linkages between the fungus and the host, or in redistribution of the fungal cell wall to facilitate branching or hyphal extension.

1.2 Endophyte bioprotective alkaloids

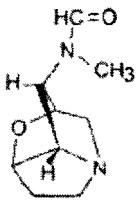
Epichloë and *Neotyphodium* endophytes produce a range of bioactive secondary metabolites (Fig. 1.2) including the ergot alkaloids (e.g. ergovaline), pyrrolizidines (e.g. lolines), a pyrrolopyrazine (e.g. peramine) and indole-diterpenes (e.g. lolitrem B). Peramine and the lolines have anti-insect properties (Rowan, Latch, 1994; Bush, et al., 1997; Wilkinson, et al., 2000). The ergot alkaloids and indole-diterpenoids are known to exhibit some protection against insects: however, they are best known for their anti-mammalian biological activity. Of the endophytes screened to date (Table 1.1), the majority were able to produce peramine, some produced ergot alkaloids and lolines and very few produced indole-diterpenes (Clay, Schardl, 2002; Siegel, et al., 1990; Christensen, et al., 1993). Many endophytes can produce at least one of these alkaloid classes, and some are able to produce more than one. Among the sexual species, *E. festucae* has the broadest alkaloid profile, with all four alkaloids reported. However, no single *E. festucae* strain has been identified that has produced all four alkaloid classes at one time (Clay, Schardl, 2002; Siegel, et al., 1990). The asexual species show greater diversity of alkaloid profiles (Table 1.1) with many isolates able to produce more than one alkaloid, consistent with their expanded genome (Section 1.1). Clay and Schardl (2002) suggest that this pyramiding of alkaloid profiles has increased the fitness of the asexual species, thereby supporting a hypothesis that the hybrids are evolving under positive selection. However, alkaloid concentrations *in planta* can be influenced by

A



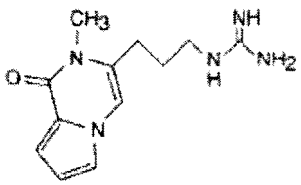
Lolitrem B

B



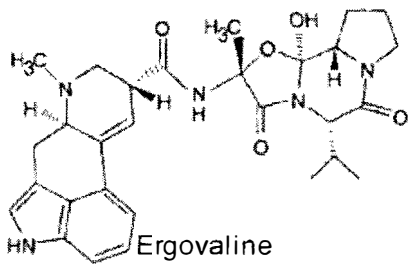
N-formylloline

C



Peramine

D



Ergovaline

Figure 1.2 Chemical structures of endophyte alkaloids

The chemical structures of (**A**) lolitrem B, (**B**) N-formylloline, (**C**) peramine and (**D**) ergovaline.

Table 1.1 Summary of alkaloid production by endophytes

Alkaloid	Producers[1]	Non-producers	Not tested
Ergot Alkaloids	*E. amarillans* *E. festucae* *N. coenophialum* *N. lolii* *Neotyphodium LpTG-2* *Neotyphodium FaTG-2* *N. inebrians (ergot alkaloids)*	*E. baconii* *E. bromicola* *E. clarkii* *E. elymi* *E. sylvatica* *E. typhina* *N. aotearoae* *N. inebrians (ergovaline)* *Neotyphodium FaTG-3* *N. occultans* *N. siegelii* *N. uncinatum*	*E. brachyelytri* *E. glyceriae* *N. australiense* *N. melicicola* *N. tembladerae*
Lolines	*E. festucae* *N. aotearoae* *N. coenophialum* *Neotyphodium FaTG-3* *N. occultans* *N. siegelii* *N. uncinatum*	*E. amarillans* *E. baconii* *E. bromicola* *E. clarkii* *E. elymi* *E. sylvatica* *E. typhina* *N. lolii* *Neotyphodium LpTG-2* *Neotyphodium FaTG-2*	*E. brachyelytri* *E. glyceriae* *N. australiense* *N. inebrians* *N. melicicola* *N. tembladerae*
Indole-diterpenes	*E. festucae* *N. lolii* *Neotyphodium FaTG-2* *(N. aotearoae)*	*E. amarillans* *E. elymi* *E. typhina* *N. coenophialum* *N. inebrians* *Neotyphodium LpTG-2* *Neotyphodium FaTG-3* *N. melicicola* *N. siegelii* *N. uncinatum*	*E. baconii* *E. brachyelytri* *E. bromicola* *E. clarkii* *E. glyceriae* *E. sylvatica* *N. australiense* *N. occultans* *N. tembladerae*
Peramine	*E. amarillans* *E. bromicola* *E. elymi* *E. festucae* *E. typhina* *N. coenophialum* *N. lolii* *Neotyphodium LpTG-2* *Neotyphodium FaTG-3* *N. tembladerae*	*E. baconii* *E. clarkii* *E. sylvatica* *Neotyphodium FaTG-2* *N. siegelii* *N. uncinatum*	*E. brachyelytri* *E. glyceriae* *N. aotearoae* *N. australiense* *N. inebrians* *N. melicicola* *N. occultans*

Data in this table have been summarised from Clay and Schardl 2002 and references there in.

[1]Producers detected in some isolates are not highlighted,
producers that are detected in most isolates are highlighted pale green,
producers that are detected in all isolates are highlighted green,
producers that have tested positive for a paxilline ELISA assay but have not had a metabolite identified are placed in parentheses.

seasonal conditions, host genotype, tissue type and age (Rowan, Latch, 1994; Siegel, Bush, 1997; Spiering, et al., 2005a).

Synthesis of these four alkaloid groups is confined to grasses infected with endophytes. This observation together with several reports of the synthesis of these alkaloids in culture confirms that ergovaline, lolines, peramine and lolitrems are of fungal origin (Penn, et al., 1993; Reinholz, Paul, 2001; Blankenship, et al., 2001). Based on these observations, one can predict that alkaloid production requires some input from the plant but whether this is through regulation, signaling and/or as a source of precursors remains to be determined.

1.2.1 Indole-diterpenes

Lolitrems (Fig. 1.2) are indole-diterpenes derived from the precursors indole and geranylgeranyl diphosphate. They are potent tremorgenic mammalian toxins known to cause ryegrass staggers in animals grazing on endophyte-infected ryegrass. Lolitrems are predominantly produced *in planta* but small quantities of indole-diterpenes have been detected in plate cultures (Penn, et al., 1993; Reinholz, Paul, 2001). Production of lolitrems is confined to the *E. festucae* and their asexual derivatives the *Neotyphodium* species (Table 1.1). Most *N. lolii* produce indole-diterpenes, but some naturally occurring isolates, such as AR1, have been identified that are devoid of lolitrems but retain many of the desirable features of endophyte-infected grass (Tapper, Latch, 1999; Fletcher, 1999).

Lolitrems have been identified in most *N. lolii* and some *E. festucae* isolates with the production of ergot alkaloids and peramine, but only one group is known to synthesise both loline and indole-diterpene alkaloids (Miles, et al., 1998). A loline-producing endophyte of *Echinopogon ovatus*, a tufted perennial grass tested positive for indole-diterpenes when examined using an ELISA-based assay with antibodies specific for paxilline analogues. Although the endophyte isolate tested positive for paxilline analogues the identity of the indole-diterpene compound produced by this strain has yet to be determined (Miles, et al., 1998). *N. coenophialum* and *N. lolii* both have an *E. festucae* ancestor, but they are only able to produce either lolines or indole-diterpenes, respectively (Bush, et al., 1997; Siegel, et al., 1990; Christensen, et al., 1993).

A biochemical pathway for lolitrem biosynthesis has been proposed from a metabolic grid based on the isolation of multiple indole-diterpene compounds from *N. lolii* infected perennial ryegrass (*Lolium perenne*) (Munday-Finch, et al., 1995; Munday-Finch, et al., 1998; Gatenby, et al., 1999; Lane, et al., 2000). However, until recently none of the genes involved in indole-diterpene biosynthesis had been isolated. The most well characterised indole-diterpene biosynthesis genes are the *pax* genes from *Penicillium paxilli*, which are required for the production of paxilline (Young, et al., 2001; McMillan, et al., 2003). Precursors of paxilline, such as paspaline and 13-desoxypaxilline, have been identified in endophyte-infected grasses suggesting they are intermediates in lolitrem production (Gallagher, et al., 1984; Weedon, Mantle, 1987; Miles, et al., 1994; Parker, Scott, 2004).

1.2.2 Ergot alkaloids

Approximately 50% of *Epichloë* and *Neotyphodium* species screened for ergot alkaloids (e.g. ergovaline, Fig. 1.2) are able to produce these compounds *in planta* (Siegel, et al., 1990). These compounds give some insect protection (Siegel, Bush, 1997) but they are best known as the agents of livestock toxicosis. Animal symptoms of ergot alkaloid toxicity include elevated body temperature, vasoconstriction and reduced weight gain.

Genes involved in ergot alkaloid biosynthesis have been cloned from *Epichloë* endophytes (Panaccione, et al., 2001; Wang, et al., 2004) and from two species of *Claviceps,* which are pathogens of cereals (Tsai, et al., 1995; Arntz, Tudzynski, 1997; Tudzynski, et al., 1999; Correia, et al., 2003). A more extensive background on ergot alkaloid biosynthesis is covered in Section 1.6.1.

1.2.3 Lolines

Lolines (Fig. 1.2) are pyrrolizidine alkaloids known for their strong anti-insect properties in endophyte-infected ryegrass. The production of lolines can exceed 10 000 μg/g of dry weight thus making them the most abundant alkaloid produced by endophytes (Bush, et al., 1993). Approximately 35% of *Epichloë*/*Neotyphodium* isolates screened were able to produce lolines (Siegel, et al., 1990 and summarised in Clay and Schardl 2002). Of the endophyte species tested, lolines were identified in *N.*

aotearoae, *N. occultans*, *N. uncinatum*, and in some *E. festucae* and *Neotyphodium* endophytes with an *E. festucae* ancestor.

Until recently lolines had only been detected *in planta*. However, recent work with a *N. uncinatum* isolate, a common endophyte of meadow fescue, resulted in the identification of conditions that produce lolines in culture (Blankenship, et al., 2001). Other grass endophytes known to produce lolines *in planta*, such as *E. festucae*, *N. coenophialum* and *N. siegelii*, failed to produce lolines under the optimal culture conditions identified for *N. uncinatum* (Blankenship, et al., 2001).

Wilkinson et al. (2000) used Mendelian genetic analysis to screen for insect feeding in a cross between Lol^+ and Lol^- *E. festucae*. The parental strains were compared to progeny with respect to their ability to produce lolines. Insect death co-segregated with certain amplified fragment length polymorphisms (AFLP) markers, which suggested the loline biosynthesis genes were at a single genetic locus, *LOL*. Subsequently, two transcripts, *lolA* and *lolC*, from a *N. uncinatum* suppression-subtractive hybridisation were identified that were up-regulated in loline-producing conditions when compared to loline-suppressed conditions (Spiering, et al., 2002b). The *lolA* and *lolC* genes are contained on a gene cluster with seven other genes proposed to be involved with loline production (Spiering, et al., 2005b). The presence of the *lol* gene cluster correlated with endophyte strains that were known to produce lolines (Spiering, et al., 2002b; Spiering, et al., 2005b).

1.2.4 Peramine

The alkaloid peramine (Fig. 1.2) acts as a feeding deterrent for *Listronotus bonariensis* (Argentine stem weevil) (Rowan, Latch, 1994). Most of the *Epichloë* and *Neotyphodium* endophytes screened to date are able to synthesise the alkaloid peramine (Table 1.1). The only *Epichloë* species that have not produced peramine are *E. baconii* and *E. clarkii* (Clay, Schardl, 2002). Peramine is a pyrrolopyrazine, that is proposed to be synthesised from proline and arginine (Lane, et al., 2000). Recently, a peptide synthetase has been cloned and characterised and shown by gene deletion to be involved in peramine production (Tanaka and Scott, unpublished results). The *perA* gene is a two-module peptide synthetase that is up-regulated *in planta*. Work is

currently underway to identify other genes that are involved in peramine production (Tanaka and Scott, personal communication).

1.3 Fungal secondary metabolism

Secondary metabolites are metabolites that are not essential for growth or are required for growth under a limited range of conditions. Such pathways are often termed dispensable. Many members of the fungal kingdom are known for the production of secondary metabolites or natural products such as antibiotics and virulence factors that have a role in fungal-host interactions. While some of these metabolites have been shown to have a role in plant virulence or are known to have anti-microbial attributes, the biological significance of many secondary metabolites is often not known.

The most well studied secondary metabolites fall into three general categories (Table 1.2), those with: (1) an isoprenoid backbone, e.g. aphidicolin, gibberellins, trichothecenes and indole-diterpenes; (2) a polypeptide backbone synthesised by a non-ribosomal peptide (NRP) synthetase, e.g. AM-toxin, ergot alkaloids, HC-toxin, penicillin and sirodesmin; or (3) a polyketide backbone synthesised by a polyketide synthase, e.g. aflatoxin, AF-toxin, AK-toxin, dothistromin, fumonisins, lovastatin, sterigmatocystin and T-toxin. However, many compounds are derivatives of combinations of these groups as well as containing other chemical modifications: e.g. ergovaline has a backbone from an NRP and an isoprenoid IPP moiety. As our knowledge of fungal systems increases other examples of secondary metabolite compounds will come to light.

1.4 Fungal gene clusters

Gene clusters are the linkages of two or more genes that participate in a common metabolic or developmental pathway (Keller, Hohn, 1997). Genes associated with clusters are considered to confer "fitness" attributes that enhance the survival of the organism (Scott, 2003). Gene clusters are well known in prokaryotes in the form of operons, which are clusters of genes with related biochemical functions driven by a

Table 1.2 Fungal secondary metabolites, their locus and structure type

Metabolite	Organism[1]	Locus[2]	Type[3]	Effect/Function[4]
Aflatoxin	*Aspergillus parasiticus, A. nidulans*	*AFLR/STC*	PKS	carcinogen
Aflatrem	*Aspergillus flavus*	*ATM*	ISO (ID)	tremorgen
AF-toxin	*Alternaria alternata*	*AFT*	PKS	pathogenicity factor
AK-toxin	*Alternaria alternata*	*AKT*	PKS	pathogenicity factor
AM-toxin	*Alternaria alternata*	*AMT*	PS	pathogenicity factor
Aphidicolin	*Phoma betae*		ISO	antiviral
Cephalosporin	*Acremonium chrysogenum*		PS	antibiotic
Compactin	*Penicillium citrinum*	*MLC*	PKS	inhibitor of HMG CoA reductase
Dothistromin	*Dothistroma pini*	*DOT*	PKS	pathogenicity factor?
Ergopeptine	*Claviceps purpurea*		PS	plant protection/anti-mammalian
Ergovaline	*Neotyphodium*		PS	plant protection/anti-mammalian
Fumonisin	*Gibberella moniliformis*	*FUM*	PKS	inhibitor of sphingolipid metabolism
Gibberellins	*Fusarium fujikuroi*		ISO	plant hormone
HC-toxin	*Cochliobolus carbonum*	*TOX2*	PS	virulence factor
Lolines	*Neotyphodium uncinatum*	*LOL*		anti-insect
Lolitrems	*Epichloë festucae, Neotyphodium lolii*	*LTM*	ISO (ID)	tremorgen
Lovastatin	*Aspergillus terrus*	*LOV*	PKS	Inhibitor of HMG CoA reductase
Paxilline	*Penicillium paxilli*	*PAX*	ISO (ID)	tremorgen
Penicillin	*Aspergillus nidulans, Penicillium chrysogenum*		PS	antibiotic
Peramine	*Epichloë festucae*	*PER*	PS	anti-insect
Sirodesmin	*Leptosphaeria maculans*	*SIR*	PS	phytotoxin
Sterigmatocystin	*Aspergillus nidulans, Penicillium chrysogenum*	*STC*	PKS	carcinogen
Trichothecenes	*Fusarium sporotrichioides, F. graminearum*	*TRI*	ISO	inhibitor of protein synthesis
T-toxin	*Cochliobolus heterostrophus*	*TOX1*	PKS	virulence factor

[1] The organisms most well characterised with respect to the secondary metabolite. Publications are cited throughout the thesis.

[2] Name of the locus/loci containing the genes involved in metabolite production.

[3] Class of compound the metabolite predominantly groups with. ISO= isoprenoid, ISO (ID)=indole diterpenoid, PKS= polyketide, PS=peptide.

[4] The known effects or the function of the metabolite.

single promoter and transcribed as a single mRNA (e.g. the *lac* operon). Fungi, which have more complex genomes than bacteria, do not appear to have operons, but instead contain gene clusters. A single promoter does not transcribe the genes in these clusters, but instead each gene is under the control of its own promoter. The close proximity of the genes in a cluster is advantageous for acquisition and maintenance of the genes within the cluster as well as facilitating the co-ordinated regulation of these genes in a similar chromatin environment.

Organisation of genes in clusters is not limited to the genes encoding enzymes for secondary metabolites. The genes for nutrient utilisation pathways, such as quinate, ethanol, proline and nitrate, are found as gene clusters in *Aspergillus nidulans* (reviewed by Keller and Hohn 1997), as are genes that determine pathogenicity in other fungi (Panaccione, et al., 1992; Han, Adams, 2001; Tanaka, et al., 1999; Johnson, et al., 2000; Pitkin, et al., 2000).

1.5 Secondary metabolites – Isoprenoids

Isoprenoids are essential for the function of a cell, as they are required for sterol production and protein prenylation. However, an increasing number of secondary metabolites have been discovered that are derived from isoprenoid precursors, such as the trichothecenes, gibberellins, indole-diterpenes and the ergot alkaloids. Isoprenoids have variable units of the five-carbon backbone, isopentenyl diphosphate (IPP). The first step in isoprenoid biosynthesis is catalysed by 3-hydroxy-3-methylglutaryl coenzyme A synthase resulting in the condensation of acetoacetyl CoA and acetyl CoA to generate 3-hydroxy-3-methylglutaryl coenzyme A (HMG CoA). The conversion of HMG CoA to mevalonate is catalysed by HMG CoA reductase in an irreversible reaction requiring NADPH/H^+. Phosphorylation and decarboxylation of mevalonate generates IPP and its isomer, dimethylallyl diphosphate (DMAPP). Geranyl diphosphate (C10), farnesyl diphosphate (C15) and geranylgeranyl diphosphate (C20) are generated by a series of C5 condensations initiated by the condensation of IPP and DMAPP in reactions catalysed by the prenyltransferases geranyl diphosphate (GPP) synthase, farnesyl diphosphate (FPP) synthase, and geranylgeranyl diphosphate (GGPP) synthase.

GGPP synthase enzymes are essential for normal growth of the fungus. GGPP is required for protein prenylation a posttranslational modification where a FPP or GGPP moiety is attached to a carboxy-terminal cysteine, which can anchor the protein to membranes (Schafer and Rine, 1992). GGPP is also a substrate in carotenoid biosynthesis (Barbato, et al., 1996). Deletions of these genes are potentially lethal or at least impair the growth of the organism. For example, inactive mutants of *al-3* (GGPP synthase) in *Neurospora crassa* (Barbato, et al., 1996) appear to be lethal, and impairment of *Saccharomyces cerevisiae* growth occurs in a *bts1* (GGPP synthetase)-negative background (Jiang, et al., 1995).

Two mechanisms for the production of secondary metabolites that possess isoprene moieties have been identified. During the production of the C15 trichothecenes, a single FPP synthase gene is up-regulated to produce sufficient C15 precursor for secondary metabolism (Tag, et al., 2001; Peplow, et al., 2003b). Fungi that synthesise diterpenes such as indole-diterpenes and gibberellins, contain two copies of a GGPP synthase gene, one of which is pathway-specific for the secondary metabolite (Young, et al., 2001; Tudzynski, Hölter, 1998; Zhang, et al., 2004).

1.5.1 Gibberellins

Gibberellins are a group of compounds known to act as plant growth regulators, regulating various stages of plant development such as stem elongation and induction of flowering. Some biosynthetic steps for gibberellin production are similar in both plant and fungi, but other steps are unique. The protein sequences of gibberellin biosynthetic enzymes such as the cytochrome P450 monooxygenases common to both *Arabidopsis* and *F. fujikuroi* have, however, very little similarity to each other (Hedden, et al., 2002). Such differences found between *Arabidopsis* and *F. fujikuroi* suggest gibberellin production has evolved independently in plants and fungi (Hedden, et al., 2002).

A cytochrome P450 monooxygenase gene involved in gibberellin biosynthesis was isolated from *F. fujikuroi* by subtractive hybridisation using RNA derived from gibberellin producing and non-producing conditions (Tudzynski, Hölter, 1998). Chromosome walking around this region revealed a cluster of seven genes that includes four P450 monooxygenases (Tudzynski, Hölter, 1998; Rojas, et al., 2001; Tudzynski, et al., 2001; Tudzynski, et al., 2002; Tudzynski, et al., 2003), *ggs2*, a geranylgeranyl

diphosphate synthase (Tudzynski, Hölter, 1998), and a copalyl diphosphate synthase (Tudzynski, Hölter, 1998). Gene replacements of P450-1, P450-2, P450-3 and P450-4 resulted in phenotypes affecting gibberellin production confirming the role of these genes in this pathway (Tudzynski, Hölter, 1998; Rojas, et al., 2001; Tudzynski, et al., 2001; Tudzynski, et al., 2003). Biochemical analysis of two gene products, P450-1 and P450-4, P450 monooxygenases, from the *F. fujikuroi* gibberellin pathway has shown that these two enzymes catalyse multiple steps (Rojas, et al., 2001; Tudzynski, et al., 2001). Recently *cpr*, a gene encoding an NADPH-dependent cytochrome P450 reductase was isolated from *F. fujikuroi*. This gene product acts as an electron donor to multiple cytochrome P450 monooxygenases and was shown to interact with the P450 monooxygenases involved in gibberellin biosynthesis as well as with other cytochrome P450 monooxygenases (Malonek, et al., 2004). The *F. fujikuroi cpr* gene is not located within the gibberellin gene cluster (Malonek, et al., 2004).

The *ggs2* was the second copy of a gene encoding a geranylgeranyl diphosphate synthetase to be isolated from *F. fujikuroi* (Tudzynski, Hölter, 1998; Mende, et al., 1997). Deletion analysis of *ggs2* confirmed it was essential for gibberellin production and was not complemented by the presence of the *ggs1* gene (Tudzynski, Hölter, 1998). This result indicates that *ggs2* is pathway specific for gibberellin production.

1.5.2 Indole diterpenes

Indole-diterpenes are compounds synthesised from an indole moiety and the C20 isoprenoid moiety geranylgeranyl diphosphate. These compounds arise from a diverse range of fungal species such as *Penicillium* (Cole, et al., 1977), *Aspergillus* (Gallagher, Wilson, 1978; Seya, et al., 1986; Nozawa, et al., 1987), *Claviceps* (Cole, et al., 1977; Dorner, et al., 1984) and *Neotyphodium* (Gallagher, et al., 1984). Many of these compounds are tremorgenic to mammals (Cole, Cox, 1981) or have biological activity against insects (Gloer, 1995; González, et al., 2003). Many indole-diterpenes, including paxilline and lolitrem B, affect calcium-activated potassium channels in smooth muscle (Knaus, et al., 1994) and are responsible for the neurotoxic disorders, such as ryegrass staggers, seen in livestock.

Until recently, our understanding of indole-diterpene biosynthesis was limited to the isolation and identification of chemical intermediates synthesised by various fungi

(Gallagher, et al., 1984; Nozawa, et al., 1987; Cole, et al., 1974; Munday-Finch, et al., 1996a; Fueki, et al., 2004). However, a 22 kb cluster of genes from *Penicillium paxilli* was isolated and shown to be involved in the biosynthesis of the indole-diterpene, paxilline (Fig. 1.3) (Young, et al., 2001; McMillan, et al., 2003; Young, et al., 1998). A core cluster of five genes *paxG, paxM, paxC, paxP* and *paxQ* encoding; a geranylgeranyl diphosphate synthase,an FAD-dependent monooxygenase, a prenyl transferase, and two cytochrome P450 monooxygenases, respectively have been shown to be essential for the production of paxilline (Young, et al., 2001; McMillan, et al., 2003). Two additional genes, *paxA* and *paxB*, have been identified within this core cluster based on significant BLASTX database matches to the available fungal genome sequences (Monahan and Scott, pers. comm).

Deletion of *paxG*, *paxM* and *paxC* results in a paxilline-negative phenotype lacking apparent intermediates (Young et al. 2001; Scott, Bryant, Astin, Saikia, Parker and Young unpublished results). Gene deletions of *paxP* and *paxQ* resulted in the accumulation of the intermediates paspaline and 13-desoxypaspaline respectively (McMillan, et al., 2003). Recent work with *Aspergillus flavus*, which produces the indole-diterpene aflatrem, a toxin similar to paxilline, identified three genes, *atmG*, *atmC* and *atmM* that have significant sequence similarity to *paxG*, *paxC* and *paxM*, respectively (Zhang, et al., 2004). This suggests that these three genes comprise an essential core of genes required for synthesis of the first stable indole-diterpene intermediate, paspaline.

The *P. paxilli* and *A. flavus* genomes encode two copies of a GGPP synthase (Young, et al., 2001; Zhang, et al., 2004). The *P. paxilli ggs1* and *A. flavus ggsB* gene products are proposed to be responsible for producing isoprenoids of primary metabolism, whereas *paxG* and *atmG* are proposed to be specific for indole diterpene biosynthesis (Young, et al., 2001; Zhang, et al., 2004). *F. fujikuroi,* a fungus that produces gibberellins (Section 1.5.1; Tudzynski and Hölter 1998) also contains two *ggs* genes. Therefore the presence of two independent *ggs* genes in a fungal genome could be a signature for diterpene production via secondary metabolism (Parker, Scott, 2004). It is expected that the aphidicolin-producing *Phoma betae* will also contain two *ggs* genes (Section 1.5.4; Toyomasu, et al., 2004).

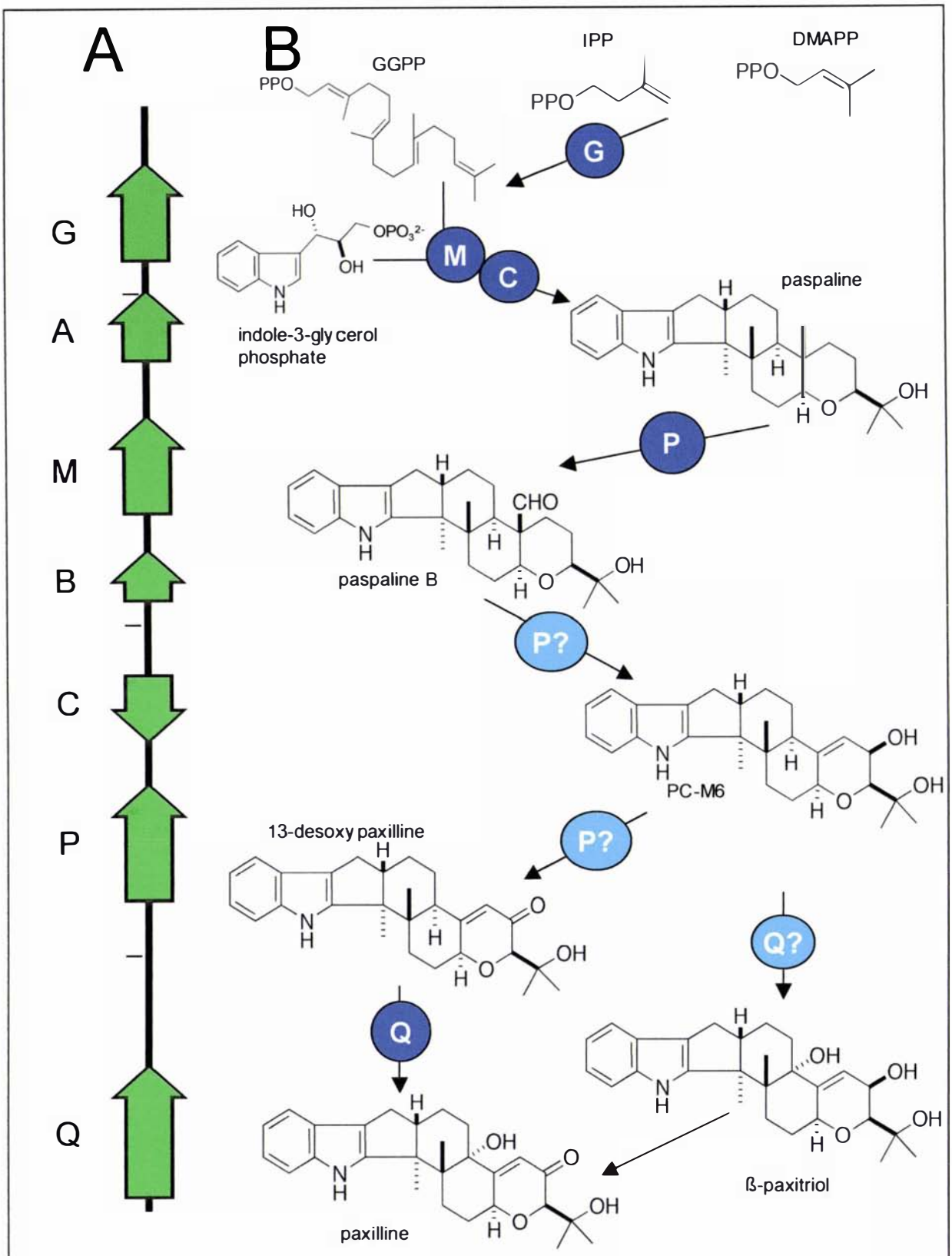


Figure 1.3 Proposed biosynthetic pathway to paxilline

The paxilline biosynthesis gene cluster from *P. paxilli* (A) and the proposed paxilline biosynthetic pathway (B). (A) The vertical line represents the cluster of core genes required for paxilline production. The gene names have been abbreviated to the single letter. (B) The proposed paxilline biosynthetic pathway. The proposed enzymatic functions based on deletion analysis of each gene is indicated by a dark blue circle. The enzymatic steps indicated in the light blue circles are putative. The functions of the gene products are discussed in Section 1.5.2.

Adjacent to the core paxilline gene cluster are two genes, *paxR* and *paxQ,* encoding putative transcription factors of the Cys_6Zn_2 zinc finger class. However, deletion of these two genes, either individually or together, had no effect on the production of paxilline (McMillan, Young and Scott, unpublished results) confirming they are not involved in paxilline regulation.

1.5.3 Trichothecenes

Trichothecenes are a diverse range of sesquiterpene (C15) compounds that share a trichodiene backbone. Trichothecenes inhibit protein synthesis in animal and human cells and are synthesised by a wide range of fungi including *Fusarium*, *Myrothecium*, *Cephalosporium*, *Trichoderma* and *Trichothecium.* Trichothecenes are proposed to function as virulence factors in maize and wheat (Harris, et al., 1999). *Fusarium sporotrichioides* and *F. subucinum* produce A-type trichothecenes that are hydroxylated at the C-8 position, whereas B-type trichothecenes contain a keto group at the C-8 position and are produced by *F. graminearum.*

The trichothecene biosynthesis genes have been extensively studied in the trichothecene-producing *F. sporotrichioides* and *F. graminearum.* Trichothecenes are produced by the cyclisation of farnesyl diphosphate by the enzyme trichodiene synthase, TRI5 (formerly known as Tox5), followed by a series of oxygenation and esterification reactions. Evidence of a gene cluster was initially shown by complementation of trichothecene mutants, *Tox4-1* and *Tox3-1*, with cosmids containing the trichodiene synthase gene (Hohn, et al., 1993). The genes involved in trichothecene production have since been shown to be contained at three co-regulated loci: (1) a major 12-gene cluster containing *Tri3* to *Tri14* (Hohn, et al., 1993; Brown, et al., 2001; Kimura, et al., 2003; Brown, et al., 2004); (2) a two-gene cluster containing *Tri1* and *Tri16* (Peplow, et al., 2003b; Meek, et al., 2003; Peplow, et al., 2003a); and (3) a single-gene locus containing *Tri101* (Kimura, et al., 1998b; McCormick, et al., 1999).

The 12-gene cluster of *F. sporotrichioides* contains seven of the 15 genes required for the biosynthetic enzymes. TRI5 is a trichodiene synthase (Hohn, et al., 1993; Hohn, Desjardins, 1992); TRI4, TRI11 and TRI13 are cytochrome P450 monooxygenases (Hohn, et al., 1995; Alexander, et al., 1998; Brown, et al., 2002; Lee, et al., 2002); TRI3 and TRI7 are acetyltransferases (Brown, et al., 2001; Kimura, et al., 2003; McCormick,

et al., 1996); and TRI8 is an esterase involved in oxygenation (Brown, et al., 2001; McCormick, Alexander, 2002). There are two regulatory genes, *Tri6* and *Tri10* (Tag, et al., 2001; Proctor, et al., 1995; Hohn, et al., 1999), an efflux pump encoded by gene *Tri12* (Alexander, et al., 1999) and *Tri9*, which has no known functional product. Three other genes contained on two separate loci encode biosynthesis genes for trichothecene production. These are *Tri101*, encoding a 3-*O*-acetyltransferase (Kimura, et al., 1998b; McCormick, et al., 1999; McCormick, et al., 1996; Kimura, et al., 1998a), and *Tri1* and *Tri16*, encoding cytochrome P450 monooxygenase and acyltransferase, respectively (Meek, et al., 2003; Peplow, et al., 2003a).

The *Tri* gene clusters of two *Fusarium* strains, *F. sporotrichioides* and *F. graminearum*, which produce the trichothecene T-2 toxin and an analogue deoxynivalenol (DON), respectively, were compared for gene order and orientation (Brown, et al., 2001). Although the *Tri* genes were found in the same order and orientation, sequence variation was identified within the intergenic regions of the two species. *Tri7*, which is required for T-2 toxin in *F. sporotrichioides* is not essential for DON synthesis in *F. graminearum* and was found to be no longer translationally functional (Brown, et al., 2001). The cytochrome P450 monooxygenase gene *Tri13* and the acyltransferase *Tri16* were also shown to be nonfunctional within the DON-producing *F. graminearum* strains (Brown, et al., 2002; Brown, et al., 2003). These results demonstrate that one mechanism for generating chemical diversity is mutation of genes in the latter steps of the pathway and channeling of those intermediates to new compounds.

Myrothecium roridum produces macrocyclic trichothecenes that are more complex than T-2 toxin. Three genes, *MRTRI4*, *MRTRI5* and *MRTRI6*, have been isolated as part of a 40 kb cluster by heterologous hybridisation of *M. roridum* DNA with the *F. sporotrichioides TRI5* gene. However the *M. roridum* genes are present in a different order and orientation to the genes from *F. sporotrichioides and F. graminearum* (Trapp, et al., 1998).

1.5.4 Aphidicolin

Aphidicolin, an inhibitor of DNA polymerase α, is a diterpene compound produced by the fungus *Phoma betae*. A gene cluster for aphidicolin biosynthesis was identified by PCR genome walking from a previously identified aphidicolin-16-β-ol synthase (ACS)

gene (Toyomasu, et al., 2004). A six-gene cluster spanning ~16 kb was identified that contained genes encoding ACS, a geranylgeranyl diphosphate synthase, two P450 monooxygenases, an ABC transporter and a putative transcription factor with a Cys_6Zn_2 motif (Toyomasu, et al., 2004).

1.6 Secondary metabolites – Non-ribosomal peptides

Secondary metabolites of the non-ribosomal peptide class have enormous structural diversity. Their synthesis is catalysed by (NRP) synthetases that have a broader specificity than the ribosome peptide synthesis machinery. NRP synthetases can use L- and D-amino acids as well as *N*-methylated amino acids, hydroxylated amino acids and many unusual amino acids. In fungi these synthetases are single multifunctional enzymes that catalyse all of the essential reactions as the enzymes have specific activation sites for each amino acid. At each activation site the specified amino acid reacts with ATP forming an aminoacyladenylate, which is noncovalently bound to the enzyme. Thioesterification covalently binds the amino acid to 4'-phosphopantetheine, releasing AMP. The first amino acid is condensed with the adjacent amino acid and so on to form a linear peptide. The peptide can then be released as a linear or cyclic product (Marahiel, et al., 1997; reviewed by Tkacz 2000). The activation sites can be highly specific for a substrate or more promiscuous and activate similar analogues leading to the synthesis of a family of compounds.

1.6.1 Ergot alkaloids

Ergopeptines, synthesised by NRP synthetases, contain lysergic acid and three amino acids that vary among the compounds of the ergopeptine family. The lysergic acid moiety is derived from the precursors tryptophan and dimethylallyl diphosphate (DMAPP). The gene encoding the determinant step in ergot alkaloid biosynthesis is *dmaW*, a dimethylally tryptophan (DMAT) synthase (Tsai, et al., 1995). Ergotamine, an abundant ergot alkaloid produced by *Claviceps purpurea,* contains the amino acids alanine, phenylalanine and proline, whereas ergovaline (Fig. 1.2) produced by *Neotyphodium lolii* contains alanine, valine and proline. Activation and coupling of the three amino acids is catalysed by a three-module non-ribosomal peptide synthetase (Panaccione, et al., 2001; Tudzynski, et al., 1999; Tudzynski, et al., 2001).

C. purpurea is a fungal pathogen of grasses that attacks young ovaries and forms ergot sclerotia. Ergopeptines are predominantly produced *in planta* at levels of ~0.5-50 μg/g of endophyte-infected grass. However, an ergopeptine de-regulated *Claviceps* strain was isolated and shown to produce ergotamine in liquid culture (Arntz, Tudzynski, 1997; Tudzynski, et al., 1999). A 22-kb gene cluster was isolated from *C. purpurea* by cross-hybridisation with *dmaW*. Contained within the cluster were genes encoding a three-module peptide synthetase (*cpps1*), two mono-module peptide synthetase (*cpps2* and *cpps3)*, a cytochrome P450 monooxygenase (*cpP450-1*) and three oxidases (*cpox1*, *cpox2* and *cpox3*) (Tudzynski, et al., 1999; Correia, et al., 2003; Tudzynski, et al., 2001). The *cpps1* gene product contains three amino acid-activating domains that are proposed to be responsible for adding the proline, phenylalanine and alanine residues to the activated lysergic acid (Tudzynski, et al., 1999). Heterologous expression of *cpps2* in *E. coli* confirms that this mono-modular non-ribosomal peptide synthetase is required for activation and probable incorporation of *D*-lysergic acid into the ergopeptine backbone (Correia, et al., 2003).

An NRP synthetase, *lpsA*, required for ergovaline biosynthesis was isolated from *N. lolii* strain Lp19, by cross hybridisation of a PCR fragment, Cp605 from *C. purpurea*, to an *N. lolii* genomic library (Panaccione, et al., 2001). The gene, *lpsA*, shares considerable identity to other peptide synthetases. The *lpsA* gene was deleted in Lp1, a *Neotyphodium* spp. LpTG-2, by homologous recombination and the single transformant isolated did not produce ergovaline *in planta* (Panaccione, et al., 2001). The *dmaW* gene from Lp1 has also been cloned and characterised. Gene disruption and complementation studies confirmed that Lp1 *dmaW* is an orthologue of the *C. purpurea* *dmaW* (Wang, et al., 2004). Gene linkage between *dmaW* and *lpsA* has not yet been established. A *cpps2* orthologue adjacent to a gene encoding an oxidase has been identified within *N. lolii* (Fleetwood, Tanaka, Scott and Johnson, pers. comm) but, again, whether or not these two genes are linked to *dmaW* and/or to *lpsA* remains to be established.

1.6.2 HC-toxin

Cochliobolus carbonum is a fungal pathogen of maize causing maize leaf spot disease. HC-toxin is a cyclic tetrapeptide and contains D-pro, L-ala, D-ala, and L-aeo (2-amino-

9, 10-epoxi-8-oxodecanoic acid), which is produced by a *C. carbonum* peptide synthetase. HC-toxin inhibits histone deacetylase, an important co-regulator of transcription. Strains that produce HC-toxin are particularly virulent on maize lines that are homozygous for the recessive allele *hm*. Maize that carries a dominant *Hm* allele can detoxify HC-toxin by the enzyme HC-toxin reductase (Meeley, et al., 1992).

The *HTS-1* gene was cloned using primers designed to the amino acid sequences of the proteins HTS-1 and HTS-2, two enzymes required for the production of HC-toxin (Scott-Craig, et al., 1992). The gene, an NRP synthetase, contains a 15.7-kb open reading frame with four adenylation motifs for activation of the four amino acids contained within HC-toxin. *C. carbonum* Tox^{+} isolates contain two functional copies of the *HTS-1* gene on a 22-kb duplicated region on the same chromosome. When both copies were inactivated the resulting transformants were unable to make HC-toxin and were non-pathogenic (Panaccione, et al., 1992). Other genes isolated as part of the HC-toxin cluster included *TOXE* encoding a regulatory gene (Ahn, Walton, 1998), *TOXA* encoding a putative HC-toxin efflux carrier for self-protection against HC-toxin (Pitkin, et al., 1996), *TOXC* encoding a fatty acid synthase β-subunit required for the addition of the side chain of Aeo (Ahn, Walton, 1997), *TOXF*, a branched-chain amino-acid transaminase, possibly required to aminate Aeo (Cheng, et al., 1999), and *TOXG* encoding an alanine racemase to synthesise D-alanine (Cheng, Walton, 2000). *HTSI* and *TOXA* (Panaccione, et al., 1992), and *TOXF* and *TOXG* (Cheng, et al., 1999) are closely linked and transcribed from divergent promoters. The *TOXE* gene product is essential for HC-toxin production and has been shown to recognise a ten-base motif, which lacks dyad symmetry, called the 'tox-box' that is present in the promoters of the *TOX2* genes (Ahn, Walton, 1998; Pedley, Walton, 2001).

The *TOX2* locus is complex as the genes are present in multiple copies covering a 600 kb region that contains other highly repetitive DNA such as transposons (Cheng, et al., 1999; Cheng, Walton, 2000; Ahn, Walton, 1996; Panaccione, et al., 1996; Ahn, et al., 2002). Some of the *TOX2* genes are tightly linked but others are dispersed across the 600 kb region (Ahn, et al., 2002). Although this region segregates as a single locus, the chromosome is unstable (Pitkin, et al., 2000).

1.6.3 AM-toxin

AM-toxin is a four-member cyclic depsipeptide produced by the *Alternaria alternata* apple pathotype. A four-module peptide synthase responsible for the production of AM-toxin was isolated from the *A. alternata* apple pathotype using primers designed to conserved regions of NRP synthetase genes (Johnson, et al., 2000). The *AMT1* gene is only detected in the apple pathotype and is located on a small, conditionally dispensable chromosome (Johnson, et al., 2001). The *AMT* locus is duplicated within the genome and contained on the same small chromosome (Johnson, et al., 2000; Akamatsu, et al., 1999). Using homologous recombination, one copy of the *AMT1* gene was disrupted in fifteen independent transformants analysed. Two of the fifteen *AMT1* deletion mutants were unable to produce AM-toxin indicating that only one copy of the gene is functional (Johnson, et al., 2000). *AMT2*, a gene encoding a protein proposed to convert 2-keto-isovaleric acid to 2-hydroxy-isovaleric acid, is contained on the same conditionally dispensable chromosome as *AMT1*, but as yet there is no evidence of direct linkage of these two genes (Ito, et al., 2004).

1.6.4 Penicillin

Penicillin is a β-lactam antibiotic produced by the filamentous fungi *Penicillium chrysogenum*, *Aspergillus nidulans*, and the cephalosporin producing *Acremonium chrysogenum*. Biosynthesis of penicillin requires an NRP synthetase to activate and condense the three amino acid precursors, L-α-aminoadipic acid, L-cysteine and L-valine. The production of penicillin requires the genes *pcbAB* encoding an ACV synthetase to produce the linear ACV molecule, and *pcbC* encoding an isopenicillin N synthase to cyclise the molecule (reviewed in Tkacz 2000). These are core genes found clustered in *P. chrysogenum*, *A. chrysogenum* and *A. nidulans* (Barredo, et al., 1989; Diez, et al., 1990; MacCabe, et al., 1990). However, *P. chrysogenum* and *A. nidulans* have a third gene in their cluster, *penDE* encoding an acyltransferase whereas *A. chrysogenum* has a second separate cluster containing *cefEF* and *cefG* genes that encode enzymes for the last two steps in the cephalosporin pathway.

1.6.5 Sirodesmin

Sirodesmin is a phytotoxin belonging to the epipolythiodioxopiperazine (ETP) class of toxins and is produced by *Leptosphaeria maculans*, a fungal pathogen of canola. ETPs are characterised by a disulphide bridge across a diketopiperazine ring. Sirodesmin is synthesised from the amino acids tyrosine and serine via an NRP synthetase, with addition of a DMAPP moiety to the tyrosine ring (Gardiner, et al., 2004). A co-regulated cluster of 18 genes with proposed biochemical functions required for sirodesmin was identified from *L. maculans*. Functional characterisation by gene disruption has confirmed that a deletion of *sirP* a gene encoding a two-module NRP synthetase, was unable to produce sirodesmin (Gardiner, et al., 2004), whereas *sirA* encoding an ABC transporter was required for autoresistance of sirodesmin (Gardiner, et al., 2005).

1.7 Secondary metabolites – Polyketides

Polyketides are a large and diverse class of natural products. Biosynthesis of polyketides occurs in a similar manner to that of long-chain fatty acids. The biosynthesis of fatty acids requires a starter unit of acetate that is condensed to an extender unit of malonate creating a β-keto group and releasing CO_2. The β-keto group is then reduced and the cycle resumes with the addition of another extender unit. Polyketides are somewhat more complex than fatty acids as they can use different extender groups and have variations in the extent of processing the β-keto group. The condensing domains are acyl transferase (AT), β-ketoacyl synthase (KS) and acyl carrier protein (ACP), and the reduction reactions require ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER). The release of the growing polyketide requires a thioesterase (TE) but cyclisation, lactonisation or amide bond formation can also occur (reviewed in Tkacz 2000). KR, DH and ER do not have to be used in polyketide formation, thus leaving the β-keto group unchanged or reduced to a hydroxyl, enoyl or alkyl function.

Multifunctional Type I polyketide synthases (PKS) are large polypeptides that encode all of the necessary enzymatic domains for one or more cycles of condensation and β-

keto processing. Type I PKS can also contain repeated modules that are clustered. All of the fungal PKS genes described below appear to be multifunctional Type I. Type II enzymes consist of several single functions or bifunctional polypeptides in a loosely associated complex (reviewed in Tkacz 2000).

1.7.1 Aflatoxin and sterigmatocystin

The polyketides aflatoxin and sterigmatocystin are derived from the same biochemical pathway. They are among the most acutely toxic, carcinogenic, mutagenic and teratogenic compounds known. They are found as natural contaminants in nuts, oilseeds and cereals and are produced by *Aspergillus flavus* (aflatoxin), *A. parasiticus* (aflatoxin) and *A. nidulans* (sterigmatocystin).

The genes required for sterigmatocystin and aflatoxin biosynthesis are found as ~60-75 kb gene clusters (Yu, et al., 1995; Brown, et al., 1996; Woloshuk, Prieto, 1998; Yu, et al., 2004; reviewed by Keller and Hohn 1997). Comparison of the aflatoxin and sterigmatocystin clusters show they possess genes encoding similar functions and regulatory control but the order of the genes is different. The clusters consist of genes encoding a polyketide synthase (Feng, Leonard, 1995; Yu, Leonard, 1995), α and β fatty acid synthases (Brown, et al., 1996), cytochrome P450 monooxygenases (Kelkar, et al., 1997; Yu, et al., 1997; Yu, et al., 1998; Keller, et al., 2000; Ehrlich, et al., 2004), dehydrogenases (Cary, et al., 1996; Chang, et al., 2000), monooxygenases (Keller, et al., 2000), ketoreductases (Chang, et al., 1993; Skory, et al., 1993; Keller, et al., 1994; Trail, et al., 1994), methyltransferases (Yu, et al., 1993) as well as regulatory proteins AFLR (Payne, et al., 1993; Woloshuk, et al., 1994) and AFLS (Meyers, et al., 1998; Chang, 2003).

The boundary of the sterigmatocystin gene cluster was determined by northern analysis as fragments that hybridised under both inducing and non-inducing conditions (Brown, et al., 1996). A gene cluster for sugar utilisation is located at the 3' end of the aflatoxin gene cluster (Yu, et al., 2000), thereby establishing a boundary of this cluster.

1.7.2 Dothistromin

Dothistromin is a polyketide that is synthesised by a similar biosynthetic pathway to aflatoxin and sterigmatocystin. Dothistromin is produced by *Dothistroma pini*, a filamentous fungus that is pathogenic to pine trees. Cross hybridisation with an *Aspergillus parasiticus ver-1* gene to a *D. pini* genomic library resulted in the isolation of a lambda clone containing: *dotA*, a *ver-1* orthologue; *dotB*, an oxidase with similarity to *stcC*; *dotC*, a toxin pump with similarity to the *C. carbonum TOXA* gene; and *dotD*, a thioesterase (Bradshaw, et al., 2002). Deletion of *dotA* eliminated dothistromin production and resulted in the accumulation of versicolorin A, a known precursor to aflatoxin and sterigmatocystin (Bradshaw, et al., 2002).

Additional genes encoding proteins with putative roles in dothistromin production were isolated using degenerate PCR with primers designed to the conserved KS domain of polyketide synthases. A partial sequence of a PKS orthologue was identified and adjacent to this gene were cytochrome P450 monooxygenase and monooxygenase genes, which are 56% and 54% identical to the sterigmatocystin genes, *stcB* and *stcW*, respectively (Morgan, and Bradshaw unpublished data). As yet there is no evidence of close linkage of the two dothistromin gene clusters.

1.7.3 AF-toxin and AK-toxin

Alternaria alternata is a plant fungal pathogen known to produce host-specific toxins. AK-toxin is the host-specific toxin produced by the Japanese pear pathotype. The AK-toxin consists of a modified polyketide precursor of which the structural backbone is also present in the AF-toxin and ACT-toxin of the strawberry and tangerine pathotypes, respectively.

Mutants in AK-toxin production were isolated using the technique of restriction enzyme-mediated integration (REMI). Isolation of the region flanking the plasmid integration identified two genes, *AKT1* and *AKT2*, a putative carboxyl-activating enzyme and a gene of unknown function respectively, which are present in the pear, strawberry and tangerine pathotypes (Tanaka, et al., 1999). Further sequence analysis of the AK-toxin locus revealed a gene, *AKTR*, encoding a transcription factor, and *AKT3* a gene encoding a putative hydratase, which are also involved in AK-toxin

production (Tanaka, Tsuge, 2000; Tanaka, Tsuge, 2001). The AK-toxin locus is contained on a 4.1-Mb chromosome and is a complex locus with tandem duplications of the *AKT* genes. However, in each case only one copy of the gene is functional (Tanaka, et al., 1999; Tanaka, Tsuge, 2000).

Homologues of the four *AKT* genes, *AKT1*, *AKT2*, *AKTR-1* and *AKT3-1*, were isolated from *A. alternata* NAF8, an AF-toxin producing strawberry pathotype. As was seen with the AK-toxin locus, the genes from NAF8 are duplicated and contained on a single 1.05-Mb conditionally dispensable chromosome. Although the sequence similarity of the four *AKT* genes to their *AFT* homologues is greater than 90% identity at the DNA level, cosmid clones from NAF8 that contained the four *AFT* genes together were never isolated (Hatta, et al., 2002). On one cosmid clone, TLS-S1, a sequence of a degenerated member of the hAT transposon family was identified. A fifth gene, *AFTS1*, with a predicted function of an aldo-ketoreductase, was identified within the *AFT* locus. *AFTS1* is present as a single copy gene on the 1.05-Mb conditionally dispensable chromosome of the strawberry pathotypes. *AMT2*, an orthologue of *AFTS1*, is present in the apple pathotype that produces AM-toxin (Section 1.6.3; Ito, et al., 2004; Hatta, et al., 2002). Ito et al. (2004) predict that the function of AFTS1 and AMT2 is to produce chemical moiety a common to both AF- and AM-toxins.

1.7.4 Fumonisins

Fumonisin, a potent inhibitor of sphingolipid biosynthesis, is produced by a number of *Gibberella* species. Fumonisins consist of a 19- or 20-carbon backbone with hydroxyl, methyl and tricarballylic acid moieties at various positions along the backbone. The backbone is very similar to that of sphinganine, a fatty acid derived intermediate in sphingolipid metabolism. However, fumonisins are polyketide derivatives synthesised by a polyketide synthase, FUM1 (Proctor, et al., 1999). FUM1 contains the conserved domains KS, AT, DH, ER, KR and ACP consistent with other polyketide synthases. Complementation of characterised fumonisin mutants, *fum1*, *fum2* and *fum3*, with cosmids that contained the gene *FUM1* showed that other genes involved in fumonisin production were contained on the cosmids. These data (Proctor, et al., 1999) supported the genetic analysis showing that the *fum1*, *fum2* and *fum3* loci in *G. moniliformis* are closely linked, suggesting a gene cluster (Desjardins, et al., 1996). Subsequently, sequence analysis of ~75-kb flanking *FUM1*, revealed 15 genes with biochemical

functions predicted to be involved in fumonisin production (Seo, et al., 2001; Proctor, et al., 2003). To date the gene cluster contains genes that encode: FUM1, a polyketide synthase; three P450 monooxygenases, FUM6, FUM12 and FUM15; FUM7, a dehydrogenase; FUM8 an aminotransferase; FUM9, a dioxygenase; FUM10 and FUM16, two fatty acyl-CoA synthetases; FUM11, a tricarboxylate transporter; FUM13, a short chain dehydrogenase; and FUM14, a gene consisting of a condensation domain common to peptide synthases. These genes are co-regulated in culture, with expression coinciding with the induction of fumonisin production (Proctor, et al., 2003). The *FUM6* gene is unusual as it is a cytochrome P450 monooxygenase fused with a cytochrome P450 reductase (Seo, et al., 2001). Deletion analysis of three genes, *FUM17*, *FUM18* and *FUM19*, adjacent to and co-regulated with the fumonisin biosynthetic genes, did not result in fumonisin-negative phenotypes, indicating their role in fumonisin production is not biosynthetic (Proctor, et al., 2003). Analyses are now underway to determine the biosynthetic pathway for fumonisin production using gene-specific deletions (Bojja, et al., 2004).

1.7.5 Lovastatin

Lovastatin is a polyketide derivative produced by *Aspergillus terrus* that is well known as an inhibitor of HMG-CoA reductase, a key enzyme involved in cholesterol biosynthesis. A gene cluster involved in lovastatin production was found to contain two polyketide synthase genes, *lovB* and *lovF*, encoding a nonaketide synthase and a diketide synthase respectively. Both polyketide synthases contain the KS, AT, DH, ER, KR and ACP domains associated with polyketide synthases, but also contain a methyltransferase domain, a feature more commonly associated with NRP synthetases (Kennedy, et al., 1999). A ~37-kb cosmid containing the polyketide synthase gene *lovB* and flanking genes was used to transform an *A. nidulans* strain devoid of lovastatin production. The resulting transformants were able to produce monacholin J, a precursor of lovastatin, and were also found to be more resistant to lovastatin. Sequence analysis of the genes contained within the transformed cosmid identified *lovC*, a putative enoyl reductase, *lovA*, a cytochrome P450 monooxygenase, *lovD*, a transesterase, and *lovE*, a transcription factor (Kennedy, et al., 1999). Functional analyses of *lovA* to *D* and *lovF* by gene disruption have shown that these genes all have a role in lovastatin production (Kennedy, et al., 1999). A gene cluster for compactin, a compound structurally similar to lovastatin, has recently been identified from *Penicillium citrinum*. The genes

identified within the *P. citrinum mlc* gene cluster have significant similarity to those of the lovastatin biosynthesis genes. Targeted disruptions of the two PKS genes, *mlcA* and *mlcB*, confirm their role in compactin biosynthesis (Abe, et al., 2002c). Introduction of additional copies of the *mlc* genes, in particular the transcription factor *mlcR*, results in enhanced production of compactin (Abe, et al., 2002a; Abe, et al., 2002b).

1.7.6 T-toxin

T-toxins are a family of long linear chain (C_{35}-C_{41}) polyketides produced by *Cochliobolus heterostrophus*, a pathogen of maize. T-toxin produced by *C. heterostrophus* race T is virulent to Texas cytoplasmic male sterile maize. T-toxin specifically binds to a protein, URF13, unique to the inner mitochondrial membrane of T-cytoplasm maize causing the formation of pores in the inner membrane and leakage of solutes necessary for normal mitochondrial function (Braun, et al., 1989). The absence of the URF13 gene results in T-toxin-insensitive mitochondria and therefore plant resistance.

The *TOX1* locus was tagged by the method of restriction enzyme mediated integration (REMI) (Lu, et al., 1994) and shown to contain a polyketide synthase, PKS1, unique to race T (Yang, et al., 1996a). PKS1 has six domains, KS, AT, DH, ER, KR and ACP. However, no thioesterase domain was found, suggesting that an additional enzyme provides this catalytic function. The *PKS1* gene contains four introns and the DNA sequences flanking *PKS1* are highly repetitive and AT rich. Race O lacks a detectable homologue, suggesting that the *PKS1* gene found in race T could have arisen by horizontal gene transfer rather than by vertical inheritance from an ancestral strain (Yang, et al., 1996a). However, recent phylogenomic analysis of fungal PKS sequences suggests that gene divergence and loss is just as likely a scenario as horizontal gene transfer (Kroken, et al., 2003).

The *TOX1* locus is in fact two loci, *Tox1A* and *Tox1B*, split between two chromosomes (Kodama, et al., 1999). The split appears to be the result of inseparable linkage between breakpoints of a reciprocal translocation of chromosomes 6 and 12 found in race T. Race T contains 1.2 Mb of unique DNA and there is evidence to show this unique DNA consists of a gene cluster for T-toxin production. The *PKS1* maps at

Tox1A, whereas a decarboxylase maps at *Tox1B* (Kodama, et al., 1999; Rose, et al., 2002).

1.8 Regulation

An obvious reason for genes encoding related biochemical functions to be clustered together, whether as an operon or a gene cluster, is to give coordinated gene regulation. Genes located in a cluster will be contained in a similar chromatin environment. However, many fungal gene clusters also contain genes encoding pathway-specific regulators within the cluster that regulate other genes for the biochemical pathway. Examples of pathway-specific regulators include AFLR and AFLS proteins required for aflatoxin and sterigmatocystin production (Chang, 2003; Yu, et al., 1996), TRI6 and TRI10 in the trichothecene cluster (Tag, et al., 2001; Proctor, et al., 1995; Hohn, et al., 1999), TOXE, which is required for the production of HC-toxin (Ahn, Walton, 1998; Pedley, Walton, 2001), MLCR, which is required for compactin production (Abe, et al., 2002a) and AKTR, which is required for AK-toxin and possibly AF-toxin production (Tanaka, Tsuge, 2000; Hatta, et al., 2002). Deletion of these genes either abolishes or greatly reduces the production of the secondary metabolite. AFLR, TRI6 and TOXE all have known binding sites to the promoter regions of the genes that comprise the cluster (Hohn, et al., 1999; Pedley, Walton, 2001; Fernandes, et al., 1998). As yet no gene encoding a transcription factor has been identified within the gibberellin or fumonisin gene clusters (Proctor, et al., 2003; Mihlan, et al., 2003). Proctor et al. (2003) have shown there are putative transcription factors adjacent to the *FUM* cluster, but the expression of these sequences is not regulated with the *FUM* genes. Putative transcription factor genes have been identified within some recently identified clusters but their involvement in regulating secondary metabolite production is yet to be confirmed (Toyomasu, et al., 2004; Gardiner, et al., 2004).

Coordinated regulation of gene clusters has been a useful feature to delineate the boundaries of gene clusters. Therefore, expression analysis has been used to distinguish genes that are involved in the production of secondary metabolites from those genes that reside adjacent to a gene cluster (Young, et al., 2001; Brown, et al., 2004; Gardiner, et al., 2004; Brown, et al., 1996; Proctor, et al., 2003; Pirttilä, et al., 2004).

Pathway-specific regulators located within the cluster do not solely regulate gene clusters. Global transcription factors such as AreA (nitrogen regulator), PacC (pH regulatory gene), FlbA, FluG (regulators of sporulation) and FadA (involved in G-protein signaling) have been shown to have an effect on the biosynthesis of the metabolites (Tudzynski, et al., 1999; Tilburn, et al., 1995; Tag, et al., 2000; Hicks, et al., 1997). Global regulators have different effects on secondary metabolite production. Aflatoxin and sterigmatocystin are produced during sporulation, a process that requires expression of FlbA and FluG (Hicks, et al., 1997; Guzmán-de-Peña, Ruiz-Herrera, 1997). However, penicillin and trichothecene production requires proliferative growth (Tag, et al., 2000). Pac1, an orthologue of PacC, is a transcriptional repressor of fumonisins so when *Pac1* was disrupted, fumonisins were produced at a higher level than from the wild-type (Flaherty, et al., 2003). On the other hand, deletion of *Zfr1*, a recently identified $Zn(II)_2Cys_6$ binuclear transcription factor from *F. verticillioides*, was shown to reduce fumonisin production by 90% (Flaherty, Woloshuk, 2004).

Recently LaeA a regulator of secondary metabolism in *Aspergillus* spp. was identified (Bok, Keller, 2004). Over-expression of LaeA increased penicillin and lovastatin production whereas deletion of *laeA* blocks expression of the genes involved in sterigmatocystin, penicillin and lovastatin biosynthesis (Bok, Keller, 2004). It remains to be seen if LaeA orthologues from fungi genetically distant to *Aspergillus* have similar effects on the expression of genes in other clusters.

In general, the genes in clusters identified to date are tightly regulated. However, an exception is the regulation of P450-3, a gibberellin biosynthesis gene that is not repressed by nitrogen and is therefore regulated differently from the remaining cluster genes (Mihlan, et al., 2003).

1.9 Autoresistance

Many metabolites produced by fungi are toxic. Sometimes they are not only toxic to the organism infected but also toxic to themselves. To overcome such toxicity, fungi have developed autoresistance mechanisms. The most commonly found autoresistance mechanism is the presence of an efflux pump. The genes for these pumps belong to the major facilitator superfamily (MFS). Efflux pumps are very efficient at protecting the

producer against its own toxic metabolite by rapidly pumping the toxin out of the cell. Efflux pumps characterised to date include TRI12 for autoprotection from the trichothecenes (Alexander, et al., 1999), TOXA for autoprotection from HC toxin (Pitkin, et al., 1996) and the *L. maculans* SIRA for autoprotection from sirodesmin (Gardiner, et al., 2005). Transporters associated with clusters that have no role in transporting the produced metabolite are SMT from the gibberellin gene cluster and AflT located within the aflatoxin cluster of *F. fujikuroi* and *A. parasiticus*, respectively (Voss, et al., 2001; Chang, et al., 2004). Putative pumps have also been identified in the dothistromin biosynthesis gene cluster (Bradshaw, et al., 2002) and the lovastatin/compactin gene clusters (Kennedy, et al., 1999; Abe, et al., 2002c), but their roles are yet to be elucidated.

Autoresistance is not limited to efflux pumps. The *P. citrinum* biosynthesis gene cluster for compactin (Section 1.5.11), an inhibitor of HMG-CoA reductase, contains a pathway-specific HMG CoA reductase, *mlcD*. Deletion of *mlcD* results in a strain that shows greater sensitivity to compactin suggesting that MlcD has a role in autoresistance (Abe, et al., 2002b). Other mechanisms for resistance include metabolism of the toxin and modification of the target site, but these mechanisms are yet to be identified in fungal secondary metabolite pathways associated with gene clusters.

1.10 Evolution of Clusters

Why are fungal genes associated with secondary metabolism clustered? Many of the clusters are comprised of genes with genetically diverse functions and therefore are thought to have arisen by unequal crossing over or gene duplication (Walton, 2000). There are no known mechanisms to maintain gene clusters but we do know there are ways of dispersing genes by translocation, inversion, and unequal crossing over. Therefore, selective pressure somehow drives and maintains clustering of these secondary metabolite genes (Walton, 2000; Rosewich, Kistler, 2000). Clustering can be rationalised as a mechanism to optimise co-regulation by *cis*-regulatory elements or by co-localisation of genes within a common chromatin environment. However, as the genes involved in housekeeping pathways are most often dispersed, this cannot be the sole requirement. Walton (2000) suggests that the evolutionary pressure that maintains the pathway as a cluster is distinct from the pressure that maintains the capacity to

produce the secondary metabolite. Transfer of a cluster into a naïve background supplies an instant selective advantage that a single gene may not, and is an effective mechanism for transfer of a biological trait. Therefore, gene clustering is advantageous for transfer of genetic material as a dispersed pathway can only be transferred vertically.

Recent phylogenomic analysis of fungal polyketide synthase genes investigated the origin and diversity of PKSs across taxonomically diverse fungal species. Phylogenomic analysis was based on the highly conserved ketosynthase domain of fungal type I PKS genes. In most cases gene duplication, divergence and/or gene loss rather than horizontal gene transfer can explain the diversity of fungal PKS genes (Kroken, et al., 2003). There were very few examples of PKSs being acquired as the result of horizontal gene transfer (Kroken, et al., 2003).

1.11 Stability of the fungal genome

1.11.1 Transposable elements

Fungal genomes contain both class I and class II transposable elements (Daboussi, 1997). The class I retroelements transpose through an RNA intermediate requiring a reverse transcriptase to undergo replicative transposition. There are two classes of retroelements, those with long terminal repeats (LTR, retrotransposons) and those without, such as long and short interspersed nuclear elements (LINEs and SINEs). Retroelements undergo a reverse trancriptase step, where the fidelity of the enzyme is low, resulting in nucleotide incorporation errors. The class II DNA transposons are copied through a DNA template and then excised, leaving a footprint of a few nucleotides at the donor site.

Many fungi display phenotypic instability, which is characteristic of organisms harbouring transposable elements. These elements can be present in varying numbers and may also be incomplete or exist only as remnants. Transposable elements have been found in the *TOX2* locus for HC-toxin production as repetitive transposons flanking the *HTS1* and *TOXA* genes (Panaccione, et al., 1996). The *PEP* gene cluster from *Nectria haematococca* contains transposon-like elements (Han, et al., 2001) and the *TOX1* locus for T-toxin has AT-rich repetitive regions (Yang, et al., 1996b). The

AFT locus required for AF-toxin production contains TLS-S1, a transposon-like sequence that belongs to the hAT transposon family (Hatta, et al., 2002).

1.11.2 Repeat-induced point mutations

Duplicated sequences within the *Neurospora crassa* genome are detected and altered when the fungus undergoes meiosis. The altering of the duplicated sequences, known as repeat induced point mutation (RIP), changes G-C pairs to A-T (Cambareri, et al., 1989). Initially thought to be exclusive to *Neurospora*, RIP has now been identified in *Podospora* (Hamann, et al., 2000; Graïa, et al., 2001) and *Leptosphaeria* (Idnurm, Howlett, 2003). Degenerate transposable elements that appear to be 'RIPed' have been identified in *Aspergillus nidulans* (Nielsen et al. 2001), *Pyricularia grisea* (Nakayashiki, et al., 1999) and *Fusarium oxysporum* (Hua-Van, et al., 1998).

1.11.3 Gene duplications and deletions

Several fungal species have duplications associated with clustered genes of secondary metabolite pathways (Tanaka, et al., 1999; Johnson, et al., 2000; Ahn, Walton, 1996; Panaccione, et al., 1996; Ito, et al., 2004; Tanaka, Tsuge, 2000; Hatta, et al., 2002; Chang, Yu, 2002). In the *Alternaria alternata* strains that produce AF-toxin, AK-toxin and AM-toxin, the genes responsible for toxin production are duplicated on the same chromosome but only one set of the duplicated genes is functional (Tanaka, et al., 1999; Johnson, et al., 2000; Ito, et al., 2004; Tanaka, Tsuge, 2000; Hatta, et al., 2002). A section of the aflatoxin gene cluster was found as a duplication in an *A. parasiticus* strain but not all the duplicated genes were functional (Chang, Yu, 2002). Multiple functional copies of the *TOX2* genes in *Cochliobolus carbonum* are located on an unstable chromosome (Pitkin, et al., 2000; Ahn, et al., 2002). Chromosome size variation has been identified with the *TOX2* genes with reciprocal chromosomal break points mapped between type 1 and type 2 strains (Ahn, Walton, 1996; Ahn, et al., 2002). The *TOX1* locus required for T-toxin production in *C. heterostrophus* maps to a single locus but is actually contained on two unlinked loci as a result of reciprocal translocation of chromosomes 6 and 12 (Kodama, et al., 1999; Rose, et al., 2002).

Duplications of genes can also result in the upregulation of the pathway. The gene cluster required for penicillin production in the high producing *Penicillium*

chrysogenum strain AS-P-78 is present in tandem repeats. Each repeat is flanked by a conserved TTTACA sequence that is also present in the natural isolates, which contain a single copy of the genes (Fierro, et al., 1993; Fierro, et al., 1995).

Deletions of secondary metabolite pathways can occur if they are contained on dispensable chromosomes (Johnson, et al., 2001; Ito, et al., 2004; Hatta, et al., 2002) or contained on large regions that are dispensable (Young, et al., 1998; Ahn, Walton, 1996; Ahn, et al., 2002). *Aspergillus oryzae* and *A. sojae* are unable to produce aflatoxin. Analysis of these strains shows that there are homologues of the aflatoxin biosynthesis genes present, but not all strains contain a complete set of pathway genes (Watson, et al., 1999; Kusumoto, et al., 2000; van den Broek, et al., 2001). Moreover, expression of the aflatoxin homologues was not detected in these *A. oryzae* or *A. sojae* isolates (Watson, et al., 1999; Kusumoto, et al., 1998).

1.12 Cloning gene clusters

Many approaches have been used to clone gene clusters for secondary metabolite biosynthesis. These include: (1) heterologous probing with an orthologue from another strain or organism (Panaccione, et al., 2001; Keller, et al., 1994; Bradshaw, et al., 2002); (2) screens using plasmid insertional mutagenesis (Young, et al., 2001; Tanaka, et al., 1999; Young, et al., 1998; Yang, et al., 1996a); (3) differential screening (Spiering, et al., 2002; Tudzynski, Hölter, 1998); (4) complementation of mutants with cosmids (Proctor, et al., 1999; Hohn, et al., 1993); (5) enzyme isolation (Scott-Craig, et al., 1992; Hohn, Beremand, 1989); (6) PCR amplification with degenerate primers to conserved domains of a proposed gene (Tsai, et al., 1995; Johnson, et al., 2000; Panaccione, et al., 1996; Proctor, et al., 1999); (7) a combined sequencing and transposon tagging system such as transposon-arrayed gene knockout (TAGKO) (Hamer, et al., 2001); (8) expression profile analysis (Pirttilä, et al., 2004; Felitti, et al., 2003); (9) whole genome sequencing (Galagan, et al., 2003; The Broad Institute). The approaches used to clone genes depend on what is known about the enzymes within the pathway, whether the proposed chemical pathway can be used to predict the type of enzymatic functions, whether similar genes have been previously isolated, whether the organism is amenable to mutation, and whether the phenotype can be easily screened for.

Putative biochemical pathways have been established for the production of indole-diterpenes based on identification of likely intermediates (Munday-Finch, et al., 1998; Munday-Finch, et al., 1996a; Mantle, Weedon, 1994). Identification of paspaline and paxilline in *N. lolii* infected *Lolium perenne* extracts suggests that these compounds are likely to be key intermediates in lolitrem production (Gatenby, et al., 1999). Support for the proposed pathway came from the cloning and characterisation of paxilline biosynthesis genes from *P. paxilli* (Fig. 1.3) (Young, et al., 2001; McMillan, et al., 2003). The paxilline biosynthesis genes are contained as a gene cluster spanning ~25-kb (Fig. 1.3). The genes in the cluster encode a GGPP synthase (*paxG*), an FAD-dependent monooxygenase (*paxM*), a prenyl transferase (*paxC*), two P450 monooxygenases (*paxP* and *paxQ*) and two genes of unknown function (*paxA* and *paxB*). Deletion of *paxG*, *paxM* and *paxC* results in mutants that are devoid of detectable paxilline analogues (Young et al. 2001; Scott, McMillan, Astin, Saikia, Young, Bryant and Parker unpublished results). PaxG is proposed to catalyse the production of GGPP the determinant step in the pathway. This is followed by the addition of indole-3-glycerol phosphate to GGPP and subsequent cyclisation catalysed by PaxM and PaxC (Parker, Scott, 2004). Deletion of *paxP* and *paxQ* results in mutants that accumulate paspaline and 13-desoxypaxilline, respectively (McMillan, et al., 2003). However, enzymes such as P450 monooxygenases could also be multifunctional, catalysing additional steps (Fig. 1.3) as seen in gibberellin biosynthesis with the P450-1 and P450-2 enzymes (Rojas, et al., 2001; Tudzynski, et al., 2001). As yet no functional roles are known for PaxA and PaxB (Monahan and Scott unpublished results).

The structural similarities of lolitrem B and paxilline, and the presence of paspaline and paxilline in endophyte-infected plant extracts (Gatenby, et al., 1999) suggest that the biosynthesis of lolitrem B is likely to initially involve similar enzymatic steps as those identified in paxilline production. Therefore, orthologues of *paxG*, *paxM*, *paxC*, *paxP* and *paxQ* would be expected within the genome of a lolitrem B producing endophyte. Isoprenylation, cyclisation and reduction are additional catalytic steps that would be required to further chemically modify the paspaline or paxilline precursors to generate lolitrem B (Fig. 1.4).

A feature of diterpene biosynthetic clusters, such as those of gibberellin, aflatrem, aphidicolin and paxilline biosynthesis, is the presence of a cluster-specific

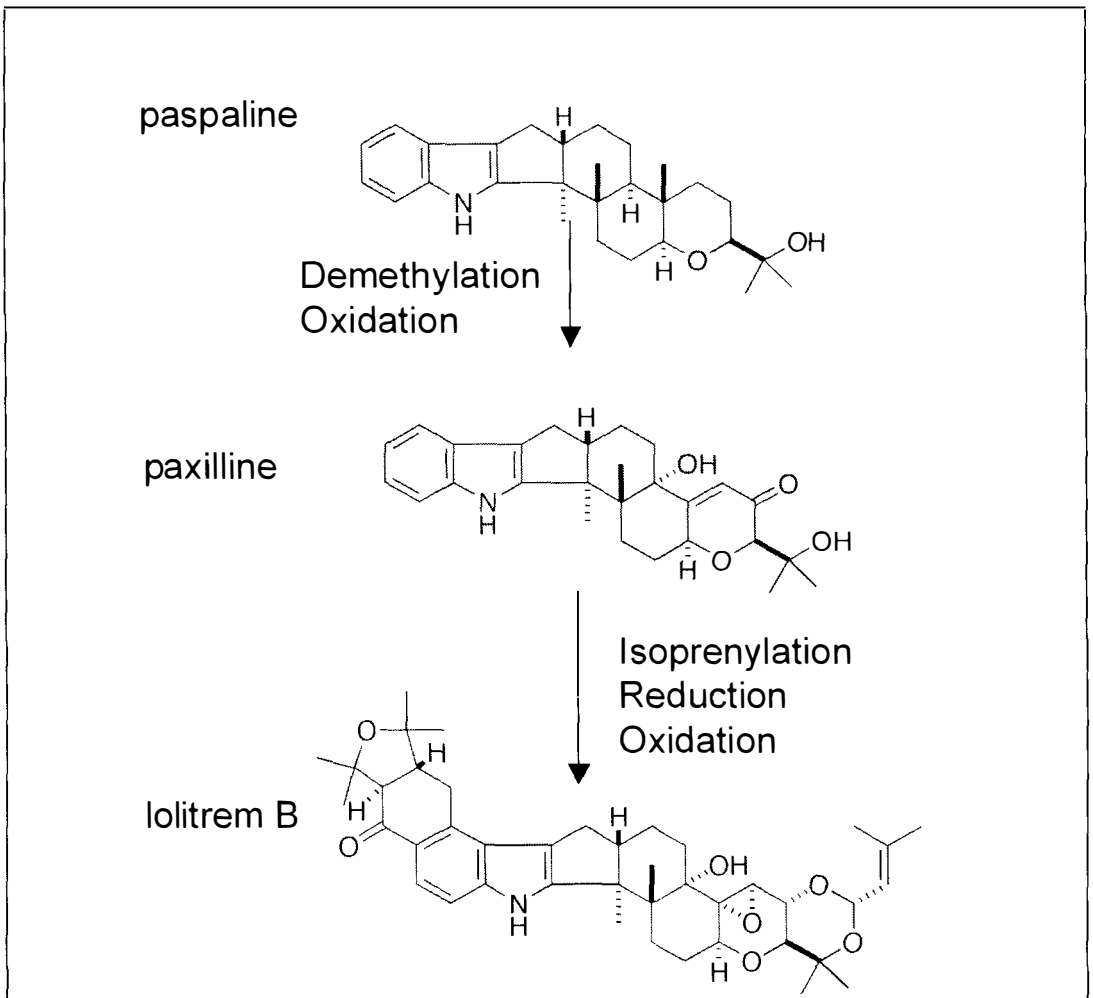


Figure 1.4 Predicted catalytic steps from paspaline to lolitrem B

Predicted modifications of paspaline to synthesise lolitrem B via paxilline.

geranylgeranyl diphosphate synthase (Young, et al., 2001; Tudzynski, Hölter, 1998; Zhang, et al., 2004; Toyomasu, et al., 2004). Furthermore, the *P. paxilli*, *F. fujikuroi* and *A. flavus* genomes all contain an additional *ggs* gene, which is presumably required for primary metabolism. Therefore, multiple copies of GGPP synthase is a useful signature of diterpene biosynthesis, and when put within the context of finding them within a gene cluster, is a useful tool for the identification of new diterpene and indole-diterpene gene clusters.

1.13 Aims

The first aim of this research is to test the hypothesis that the *Epichloë* endophytes that have the capacity to synthesise lolitrems will contain two copies of GGPP synthase, one of which will be clustered with additional lolitrem biosynthesis genes. The lolitrem biosynthesis genes are then predicted to be contained as a gene cluster that will be co-regulated and preferentially expressed *in planta*. This gene cluster is predicted to only be contained in the *Epichloë* endophytes that have the ability to produce indole diterpenes such as lolitrem B.

Based on the above hypotheses this project has been separated into four key objectives:
(1) to clone and characterise the lolitrem B biosynthesis genes;
(2) determine the expression profiles of these genes;
(3) functionally characterise at least one of these genes by deletion analysis and complementation of mutants from the *P. paxilli* paxilline biosynthesis pathway;
(4) establish the taxonomic distribution of the lolitrem biosynthesis genes and determine whether this is reflected in the indole diterpene chemotype of the isolates that were analysed.

In objective one, degenerate PCR was used to isolate the two GGPP synthase genes from the lolitrem B producing isolates *N. lolii* Lp19 and *E. festucae* Fl1. One of these gene products is expected to be required for lolitrem biosynthesis and is predicted to only be present in the lolitrem producing isolates. A least 10 - 12 genes will be required for the production of lolitrems and these genes are predicted to be present in a gene cluster at a single locus. Each identified gene was analysed for the intron/exon structures using RT-PCR and DNA sequence analysis. Expression profiles of each gene

were used to determine when the genes are expressed and whether these expression patterns are consistent with the observation that lolitrems are predominatly produced in endophyte-infected ryegrass.

One lolitrem biosynthesis gene was selected for functional characterisation by deletion analysis to verify an involvement in the biosynthesis of indole diterpenes. Complementation of the *P. paxilli pax* deletion mutants was used to confirm genes that are predicted to be orthologues of the *pax* genes.

The final objective addressed the taxonomic distribution of the lolitrem biosynthesis genes to establish why so few *Epichloë* and *Neotyphodium* endophytes are unable to produce indole diterpene compounds.

CHAPTER TWO

MATERIALS AND METHODS

2.1 Biological material

Fungal and bacterial strains, plasmids and λ clones used throughout this study are listed in Table 2.1.

2.2 Growth of cultures

All media were prepared with MilliQ water and sterilised at 121°C for 15 min, unless otherwise stated.

2.2.1 Aspergillus Complete Media (ACM)

ACM contained 2% (w/v) malt extract, 1% (w/v) Bacto-peptone, 2% (w/v) glucose, 1.5% (w/v) agar.

2.2.2 CD + Yeast Extract Media (CDYE)

CDYE media contained (per L) 3.34% (w/v) Czapek Dox (CD, Oxoid), 0.5% yeast extract and 5 mL of trace elements (Section 2.2.6).

2.2.3 LB Media

LB media (Miller, 1972) contained 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl. The pH was adjusted to 7.0 prior to autoclaving. LB agar was made by the addition of agar to a final concentration of 1.5% (w/v). LB agarose was made by the addition of agarose to a final concentration of 1.5% (w/v) agarose.

2.2.4 Media Supplements

Where appropriate the media was supplemented as follows: ampicillin, 100 μg/mL; geneticin, 150 and 200 μg/mL; hygromycin, 100, 150 and 200 μg/mL; IPTG, 38 μg/mL; tetracycline, 15 μg/mL; X-gal, 32 μg/mL.

2.2.5 Potato Dextrose Media

PD broth contains 2.4% (w/v) potato dextrose broth (Difco). PD agar was prepared by addition of agar to a final concentration of 1.5% (w/v).

Table 2.1 Biological material

Biological material	strain or identifier	Relevant characteristics	Reference
Fungal strains			
Epichloe amarillans	E52	Host *Sphenopholis obtusata*	ATCC 200743, Schardl et al, 1997
	E57	Host *Agrostis hiemalis*	ATCC 200744, Schardl et al, 1997
E. baconii	E248	Host *Agrostis stolonifera*	ATCC 765527
	E1031	Host *Calamagrostis villosa*	Moon and Schardl, Schardl et al, 1997
E. brachyelytri	E1040	Host *Brachyelytrum erectum*	ATCC 200752, Schardl and Leuchtmann, 1999
E. bromicola	E501	Host *Bromus erectus*	ATCC 200749, Leuchtmann and Schardl, 1998
	E799	Host *Bromus benekenii*	ATCC 201559, Leuchtmann and Schardl, 1998
E. clarkii	E422	Host *Holcus lanatus*	ATCC 90768, Schardl et al, 1997
E. elymi	E56	Host *Elymus canadensis*	ATCC 201551, Schardl and Leuchtmann, 1999
	E184	Host *Elymus virginicus*	ATCC 200850, Schardl and Leuchtmann, 1999
E. festucae			
PN2130	Fr1	Host *Festuca rubra*	Leuchtmann, 1994
PN2131	Frc5	Host *Festuca rubra* ssp. commutata	M. Christensen (AgResearch, NZ)
PN2133	Frr1	Host *Festuca rubra* ssp. rubra	M. Christensen (AgResearch, NZ)
PN2132	Frc7	Host *Festuca rubra* ssp. commutata	M. Christensen (AgResearch, NZ)
PN2134	Fg1	Host *Festuca glauca*	M. Christensen (AgResearch, NZ)
PN2241	E189	Host *Festuca rubra* ssp. rubra	Schardl et al, 1994
PN2291	Fl1	Host *Festuca longifolia*; wild-type; Lolitrem B positive	Leuchtmann, 1994
PN2304	CYFl1-M12	PN2291/Δ*ltmM*::P*trpC-hph*; Hyg^R	This study
PN2303	CYFl1-M28	PN2291/Δ*ltmM*::P*trpC-hph*; Hyg^R; lolitrem negative	This study
PN2302	CYFl1-M43	PN2291/Δ*ltmM*::P*trpC-hph*; Hyg^R	This study
PN2301	CYFl1-M61	PN2291/Δ*ltmMG*::P*trpC-hph*; Hyg^R; lolitrem negative	This study
PN2300	CYFl1-M65	PN2291/Δ*ltmM*::P*trpC-hph*; Hyg^R	This study
PN2299	CYFl1-M80	PN2291/Δ*ltmM*::P*trpC-hph*; Hyg^R	This study
PN2298	CYFl1-M98	PN2291/Δ*ltmM*::P*trpC-hph*; Hyg^R	This study
PN2297	CYFl1-M101	PN2291/Δ*ltmM*::P*trpC-hph*; Hyg^R	This study
PN2296	CYFl1-M142	PN2291/Δ*ltmM*::P*trpC-hph*; Hyg^R	This study
PN2295	CYFl1-M145	PN2291/Δ*ltmM*::P*trpC-hph*; Hyg^R	This study
PN2294	CYFl1-M151	PN2291/*ltmM*5'-P*trpC-hph-ltmM*3'; Hyg^R; ectopic integration; lolitrem positive	This study
PN2293	CYFl1-M156	PN2291/Δ*ltmM*::P*trpC-hph*; Hyg^R	This study
PN2292	CYFl1-M160	PN2291/*ltmM*5'-P*trpC-hph-ltmM*3'; Hyg^R; ectopic integration; lolitrem positive	This study
PN2343	2301-ltmM3	PN2301/pCY40; Hyg^R Gen^R; lolitrem negative	This study
PN2342	2301-ltmM6	PN2301/pCY40; Hyg^R Gen^R; lolitrem negative	This study
PN2344	2301-ltmM7	PN2301/pCY40; Hyg^R Gen^R; lolitrem negative	This study
PN2345	2301-ltmM10	PN2301/pCY40; Hyg^R Gen^R; lolitrem negative	This study
PN2348	2301-pll99-4	PN2301/pll99; Hyg^R Gen^R; lolitrem negative	This study
PN2349	2301-pll99-10	PN2301/pll99; Hyg^R Gen^R; lolitrem negative	This study
PN2338	2303-ltmM2	PN2303/pCY40; Hyg^R Gen^R; lolitrem positive	This study
PN2339	2303-ltmM3	PN2303/pCY40; Hyg^R Gen^R; lolitrem positive	This study
PN2340	2303-ltmM9	PN2303/pCY40; Hyg^R Gen^R; lolitrem positive	This study
PN2341	2303-ltmM10	PN2303/pCY40; Hyg^R Gen^R; lolitrem positive	This study
PN2350	2303-paxM2	PN2303/pCY41; Hyg^R Gen^R; lolitrem negative	This study
PN2351	2303-paxM4	PN2303/pCY41; Hyg^R Gen^R; lolitrem negative	This study
PN2352	2303-paxM7	PN2303/pCY41; Hyg^R Gen^R; lolitrem positive	This study
PN2346	2303-pll99-6	PN2303/pll99; Hyg^R Gen^R; lolitrem negative	This study
PN2347	2303-pll99-8	PN2303/pll99; Hyg^R Gen^R; lolitrem negative	This study
E. glyceriae	E2772	Host *Glyceria striata*	ATCC 200755, Schardl and Leuchtmann, 1999
	E277	Host *Glyceria striata*	ATCC 200747, Schardl and Leuchtmann, 1999
E. sylvatica	E503	Host *Brachypodium sylvaticum*	ATCC 200751, Leuchtmann and Schardl, 1998
	E354	Host *Brachypodium sylvaticum*	ATCC 200748, Leuchtmann and Schardl, 1998
E. typhina			
PN2238	E8	Host perennial ryegrass; lolitrem B negative	Schardl et al, 1994
	E505	Host *Brachypodium pinnatum*	ATCC 200739, Schardl et al, 1997
	E425	Host *Phleum pratense*	ATCC 200851, Schardl et al, 1997
	E1022	Host *Poa nemoralis*	ATCC 201668, Leuchtmann and Schardl, 1998
	E348	Host *Phleum pratense*	CBS 102648
	E2463	Host *Dactylis glomerata*	Moon and Schardl, University of Kentucky
	Poa	Lolitrem B negative	M. Christensen, AgResearch, NZ
Neotyphodium aotearoae	E899	Host *Echinopogon ovatus*	Moon and Schardl, University of Kentucky
N. australiense	E938	Host *Echinopogon ovatus*	Moon and Schardl, University of Kentucky
N. coenophialum			
PN2149	Tf28	Host *Lolium arundinaceum*	Christensen et al, 1993
N. huerfanum	E4055	Host *Festuca arizonica*	Moon and Schardl, University of Kentucky
N. inebrians	E818	Host *Achnatherum inebrians*	Moon and Schardl, University of Kentucky
N. lolii			
PN2191	Lp19	Host *Lolium perenne*; wild-type; lolitrem B positive	Christensen et al, 1993
PN2123	Lp5	Host *Lolium perenne*	Christensen et al, 1993
PN2136	Lp7	Host *Lolium perenne*	Christensen et al, 1993
PN2138	Lp14 (AR37)	Host *Lolium perenne*; janthitrem positive	Christensen et al, 1993
PN2141	AR1	Host *Lolium perenne*; lolitrem B negative	M. Christensen (AgResearch, NZ)
N. melicicola	E822	Host *Melica racemosa*	Moon and Schardl, University of Kentucky
N. siegelii	E915	Host *Lolium pratense*	ATCC 74483
N. sp. FaTG-2			
PN2128	Tf15	Host *Lolium arundinaceum*	Christensen et al, 1993
PN2147	Tf20	Host *Lolium arundinaceum*	Christensen et al, 1993
N. sp. LpTG-2			
PN2197	Lp1	Host *Lolium perenne*; lolitrem B negative	Christensen et al, 1993
PN2122	Lp2	Host *Lolium perenne*	Christensen et al, 1993
N. sp. 'sleepygrass'	E4096	Host *Stipa robusta*	Moon and Schardl, University of Kentucky
N. tembladarae	E1169	Host *Poa huecu*	ATCC 200844
N. sp.			
PN2126	Hd1	Host *Hordeum bogdanii*	M. Christensen (AgResearch, NZ)
Penicillium paxilli			
PN2013		Wild-type; paxiline positive	Itoh et al, 1994
PN2253	LM662	PN2013/ΔpaxT-paxD; Hyg^R; paxilline negative	Young et al, 2001
PN2257, PN2375	LMM100	PN2013/ΔpaxM; Hyg^R; paxilline negative	McMillan and Scott, unpublished
PN2290, PN2353	ABC83	PN2013/ΔpaxC; Hyg^R; paxilline negative	Bryant, Astin and Scott, unpublished
PN2376	LMM200-1	PN2257/pll99; Hyg^R Gen^R; paxilline negative	This study
PN2377	LMM200-2	PN2257/pll99; Hyg^R Gen^R; paxilline negative	This study
PN2378	LMM300-1	PN2257/pCY40; Hyg^R Gen^R; paxilline positive	This study
PN2379	LMM300-2	PN2257/pCY40; Hyg^R Gen^R; paxilline positive	This study

Biological material	strain or identifier	Relevant characteristics	Reference
PN2380	LMM300-3	PN2257/pCY40; Hyq^{P} Gen^{R}; paxilline positive	This study
PN2381	LMM300-4	PN2257/pCY40; Hyq^{P} Gen^{R}; paxilline positive	This study
PN2382	LMM400-1	PN2257/pCY41; Hyq^{P} Gen^{R}; paxilline negative	This study
PN2383	LMM400-2	PN2257/pCY41; Hyq^{P} Gen^{R}; paxilline negative	This study
PN2384	LMM400-3	PN2257/pCY41; Hyq^{P} Gen^{R}; paxilline positive	This study
PN2385	LMM400-4	PN2257/pCY41; Hyq^{P} Gen^{R}; paxilline positive	This study
PN2386	LMM400-5	PN2257/pCY41; Hyq^{P} Gen^{R}; paxilline negative	This study
PN2387	LMM500-1	PN2257/pCY54/pII99; Hyq^{P} Gen^{R}; paxilline positive	This study
PN2388	LMM500-2	PN2257/pCY54/pII99; Hyq^{P} Gen^{R}; paxilline negative	This study
PN2389	LMM500-4	PN2257/pCY54/pII99; Hyq^{P} Gen^{R}; paxilline negative	This study
PN2390	LMM500-7	PN2257/pCY54/pII99; Hyq^{P} Gen^{R}; paxilline positive	This study
PN2391	LMM500-8	PN2257/pCY54/pII99; Hyq^{P} Gen^{R}; paxilline positive	This study
PN2392	LMM500-9	PN2257/pCY54/pII99; Hyq^{P} Gen^{R}; paxilline negative	This study
PN2393	LMM600-2	PN2257/pCY55/pII99; Hyq^{P} Gen^{R}; paxilline positive	This study
PN2394	LMM600-3	PN2257/pCY55/pII99; Hyq^{P} Gen^{R}; paxilline positive	This study
PN2395	LMM600-4	PN2257/pCY55/pII99; Hyq^{P} Gen^{R}; paxilline positive	This study
PN2396	LMM600-5	PN2257/pCY55/pII99; Hyq^{P} Gen^{R}; paxilline negative	This study
PN2397	LMM600-6	PN2257/pCY55/pII99; Hyq^{P} Gen^{R}; paxilline positive	This study
PN2398	LMM600-7	PN2257/pCY55/pII99; Hyq^{P} Gen^{R}; paxilline negative	This study
PN2399	LMM600-8	PN2257/pCY55/pII99; Hyq^{P} Gen^{R}; paxilline positive	This study
PN2400	LMM600-9	PN2257/pCY55/pII99; Hyq^{P} Gen^{R}; paxilline negative	This study
PN2401	LMM600-10	PN2257/pCY55/pII99; Hyq^{P} Gen^{R}; paxilline positive	This study
PN2354	ABC283-1	PN2290/pII99; Hyq^{R} Gen^{R}; paxilline negative	This study
PN2355	ABC283-2	PN2290/pII99; Hyq^{R} Gen^{R}; paxilline negative	This study
PN2356	ABC383-1	PN2290/pCY66/pII99; Hyq^{R} Gen^{R}; paxilline negative	This study
PN2357	ABC383-2	PN2290/pCY66/pII99; Hyq^{R} Gen^{R}; paxilline negative	This study
PN2358	ABC383-3	PN2290/pCY66/pII99; Hyq^{R} Gen^{R}; paxilline negative	This study
PN2359	ABC383-4	PN2290/pCY66/pII99; Hyq^{R} Gen^{R}; paxilline negative	This study
PN2416	ABC383-5	PN2290/pCY66/pII99; Hyq^{R} Gen^{R}; paxilline negative	This study
PN2360	ABC483-1	PN2290/pJA8; Hyq^{R} Gen^{R}; paxilline positive	This study
PN2361	ABC483-2	PN2290/pJA8; Hyq^{R} Gen^{R}; paxilline positive	This study
PN2362	ABC483-3	PN2290/pJA8; Hyq^{R} Gen^{R}; paxilline positive	This study
PN2363	ABC483-4	PN2290/pJA8; Hyq^{R} Gen^{R}; paxilline positive	This study
PN2364	ABC483-5	PN2290/pJA8; Hyq^{R} Gen^{R}; paxilline positive	This study
PN2365	ABC583-1	PN2290/pCY34/pII99; Hyq^{R} Gen^{R}; paxilline negative	This study
PN2366	ABC583-2	PN2290/pCY34/pII99; Hyq^{R} Gen^{R}; paxilline positive	This study
PN2367	ABC583-3	PN2290/pCY34/pII99; Hyq^{R} Gen^{R}; paxilline positive	This study
PN2368	ABC583-4	PN2290/pCY34/pII99; Hyq^{R} Gen^{R}; paxilline positive	This study
PN2369	ABC583-5	PN2290/pCY34/pII99; Hyq^{R} Gen^{R}; paxilline negative	This study
PN2370	ABC583-6	PN2290/pCY34/pII99; Hyq^{R} Gen^{R}; paxilline positive	This study
PN2371	ABC583-7	PN2290/pCY34/pII99; Hyq^{R} Gen^{R}; paxilline positive	This study
PN2372	ABC583-8	PN2290/pCY34/pII99; Hyq^{R} Gen^{R}; paxilline positive	This study
PN2373	ABC583-9	PN2290/pCY34/pII99; Hyq^{R} Gen^{R}; paxilline positive	This study
PN2374	ABC583-10	PN2290/pCY34/pII99; Hyq^{R} Gen^{R}; paxilline positive	This study
Plant material			
Lolium perenne	Contains		
G1112	CYFI-M28	*L. perenne*/PN2303	This study
G1113	CYFI-M28	*L. perenne*/PN2303	This study
G1114	CYFI-M28	*L. perenne*/PN2303	This study
G1115		*L. perenne*; endophyte free	This study
G1116	CYFI-M28	*L. perenne*/PN2303	This study
G1117	CYFI-M28	*L. perenne*/PN2303	This study
G1118	CYFI-M61	*L. perenne*/PN2301	This study
G1119	CYFI-M61	*L. perenne*/PN2301	This study
G1120	CYFI-M61	*L. perenne*/PN2301	This study
G1121		*L. perenne*; endophyte free	This study
G1122	CYFI-M61	*L. perenne*/PN2301	This study
G1123	CYFI-M142	*L. perenne*/PN2296	This study
G1124	CYFI-M142	*L. perenne*/PN2296	This study
G1125	CYFI-M142	*L. perenne*/PN2296	This study
G1126	CYFI-M142	*L. perenne*/PN2296	This study
G1127	CYFI-M142	*L. perenne*/PN2296	This study
G1128	CYFI-M151	*L. perenne*/PN2294	This study
G1129	CYFI-M151	*L. perenne*/PN2294	This study
G1130	CYFI-M151	*L. perenne*/PN2294	This study
G1131	CYFI-M151	*L. perenne*/PN2294	This study
G1132	CYFI-M151	*L. perenne*/PN2294	This study
G1133	FI1	*L. perenne*/PN2291	This study
G1135	FI1	*L. perenne*/PN2291	This study
G1136	FI1	*L. perenne*/PN2291	This study
G1137	FI1	*L. perenne*/PN2291	This study
G1138		*L. perenne*; endophyte free	This study
G1167	2303-ltmM2	*L. perenne*/PN2338	This study
G1168	2303-ltmM2	*L. perenne*/PN2338	This study
G1169	2303-ltmM3	*L. perenne*/PN2339	This study
G1170	2303-ltmM3	*L. perenne*/PN2339	This study
G1171	2303-ltmM3	*L. perenne*/PN2339	This study
G1172	2303-ltmM9	*L. perenne*/PN2340	This study
G1173	2303-ltmM9	*L. perenne*/PN2340	This study
G1174	2303-ltmM10	*L. perenne*/PN2341	This study
G1175	2303-ltmM10	*L. perenne*/PN2341	This study
G1176	2303-pII99-6	*L. perenne*/PN2346	This study
G1177	2303-pII99-6	*L. perenne*/PN2346	This study
G1178	2303-pII99-8	*L. perenne*/PN2347	This study
G1179	2303-pII99-8	*L. perenne*/PN2347	This study
G1180	2303-paxM2	*L. perenne*/PN2350	This study
G1181	2303-paxM2	*L. perenne*/PN2350	This study
G1182	2303-paxM4	*L. perenne*/PN2351	This study
G1183	2303-paxM4	*L. perenne*/PN2351	This study
G1184	2303-paxM7	*L. perenne*/PN2352	This study
G1185	2303-paxM7	*L. perenne*/PN2352	This study
G1186	2301-ltmM3	*L. perenne*/PN2343	This study
G1187	2301-ltmM3	*L. perenne*/PN2343	This study
G1188	2301-ltmM6	*L. perenne*/PN2342	This study
G1189	2301-ltmM6	*L. perenne*/PN2342	This study

Biological material	strain or identifier	Relevant characteristics	Reference
G1190	2301-ltmM7	*L. perenne*/PN2344	This study
G1191	2301-ltmM7	*L. perenne*/PN2344	This study
G1192	2301-ltmM10	*L. perenne*/PN2345	This study
G1193	2301-ltmM10	*L. perenne*/PN2345	This study
G1194	2301-pll99-4	*L. perenne*/PN2348	This study
G1195	2301-pll99-4	*L. perenne*/PN2348	This study
G1196	2301-pll99-10	*L. perenne*/PN2349	This study
G1197	2301-pll99-10	*L. perenne*/PN2349	This study
G1198	Fl1	*L. perenne*/PN2291	This study
G1199	Fl1	*L. perenne*/PN2291	This study
G1200	CYFl1-M28	*L. perenne*/PN2303	This study
G1201	CYFl1-M28	*L. perenne*/PN2303	This study
G1202	CYFl1-M61	*L. perenne*/PN2301	This study
G1203	CYFl1-M61	*L. perenne*/PN2301	This study
G1204		*L. perenne*; endophyte free	This study
G1205		*L. perenne*; endophyte free	This study
Bacterial strains			
XL1-Blue		*supE44 hsdR17 recA1 gyrA46 thi relA1 Lac F'[proAB$^+$ lacI^q lacZΔM15 Tn10 (Tet^r)]*	Bullock et al, 1987
KW251		F *supE44*galT22 metB1 *hsdR2* mcrA (argA81:Tn10] *recD1014 Tet^s*	Promega
PN1676		XL1/pCY28	This study
PN1677		XL1/pCY29	This study
PN1855		XL1/pCY34	This study
PN1856		XL1/pCY39	This study
PN1858		XL1/pCY40	This study
PN1857		XL1/pCY41	This study
PN1853		XL1/pCY54	This study
PN1854		XL1/pCY55	This study
PN1859		XL1/pCY66	This study
PN1688		XL1/pPN1688	This study
PN1851		XL1/pPN1851	This study
Plasmids			
pCB1004		Amp^R/Hyg^R (P*trpC-hph*)	Carroll et al, 1994
pCY29		pGEM-T containing a 272-bp *qgs1* PCR fragment; Amp^R	This study
pCY28		pGEM-T containing a 209-bp *ltmG* PCR fragment; Amp^R	This study
pCY34		pPN1851 containing a *ltmC* 1.2 kb *NcoI/EcoR*I fragment; Amp^R	This study
pCY39		pUC118 containing *NltmM*5'-P*trpC-hph-ltmM*3'; Amp^R/Hyg^R	This study
pCY40		pll99 containing *NltmM* on a 7 kb *Xho*I fragment; Amp^R/Gen^R	This study
pCY41		pll99 containing *PpaxM* on a 3.7 kb *Bam*I fragment; Amp^R/Gen^R	This study
pCY47		pGEM-T easy containing a *ltmG* cDNA amplified with primers lol79 and lol157	This study
pCY48		pGEM-T easy containing a *ltmM* cDNA amplified with primers lol7 and lol28	This study
pCY49		pGEM-T easy containing a *ltmk* cDNA amplified with primers lol32 and lol8	This study
pCY54		pPN1851 containing *NltmM* as a 1.8 kb *NcoI/Sst*I fragment; Amp^R	This study
pCY55		pPN1851 containing *EltmM* as a 1.8 kb *NcoI/Sst*I fragment; Amp^R	This study
pCY66		pUC118 containing a 3.5 kb *Hin*dIII *ltmC* fragment; Amp^R	This study
pGEM-T		Amp^R	Promega
pGEM-T easy		Amp^R	Promega
pll99		Amp^R/Gen^R (P*trpC-nptII*-T*trpC*)	Namiki et al, 2001
pJA8		Amp^R/Hyg^R (P*trpC-hph*)	Bryant, Astin and Scott, unpublished
pPN1688	pCYhph	Amp^R/Hyg^R (P*trpC-hph*)	This study
pPN1851		pUC19 containing *Bam*HI fragment of 850 bp of *paxM* promoter amplified with primers CY9 and CY11; Amp^R	This study
pUC19		Amp^R	Messing 1983
pUC118		Amp^R	Vieira and Messing, 1987
Lambda clones			
λCY46		λGEM-11 clone from *P. paxilli* containing *paxM*	Young et al, 2001
λCY100		λGEM-12 clone from *N. lolii* containing *ggs1*	This study
λCY102		λGEM-12 clone from *N. lolii* containing *ggs1*	This study
λCY218		λGEM-12 clone from *N. lolii* containing *ltmG*	This study
λCY219		λGEM-12 clone from *N. lolii* containing *ltmG*	This study
λCY224		λGEM-12 clone from *N. lolii* containing *Rua*	This study
λCY225		λGEM-12 clone from *N. lolii* containing *Rua*	This study
λCY226		λGEM-12 clone from *N. lolii* containing *Rua*	This study
λCY227		λGEM-12 clone from *N. lolii* containing *Rua*	This study
λCY229		λGEM-12 clone from *N. lolii* containing *Rua*	This study
λCY231		λGEM-12 clone from *N. lolii* containing *Rua*	This study
λCY255		λGEM-12 clone from *N. lolii* containing *ltmK*	This study
λCY275		λGEM-12 clone from *N. lolii* containing *pks*	This study
λCY274		λGEM-12 clone from *N. lolii* containing *pks*	This study
λCY300		λGEM-12 clone from *N. lolii* containing *ltmC*	This study
λCY301		λGEM-12 clone from *N. lolii* containing *ltmC-ltmP*	This study
λCY305		λGEM-12 clone from *N. lolii* containing *ltmC*	This study
λCY324		λGEM-12 clone from *N. lolii* containing *ltmJ*	This study
λCY325		λGEM-12 clone from *N. lolii* containing *ltmJ* and *chsV*	This study
λCY338		λGEM-12 clone from *N. lolii* containing *ltmJ* and *chsV*	This study
λCY344		λGEM-12 clone from *N. lolii* containing *ltmC-ltmP*	This study
λCY346		λGEM-12 clone from *N. lolii* containing *ltmP*	This study

Genomic DNA from all ATCC, CBS or strains was supplied by Moon and Schardl (University of Kentucky).

2.2.6 SOC Media

SOC media (Dower, et al., 1988) contained: 20 mM glucose, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4.7H_2O$, 10 mM NaCl, 2% (w/v) tryptone, and 0.5% (w/v) yeast extract.

2.2.7 Trace Elements

Trace Element Mix contained (g/L): 0.5 $FeSO_4.7H_2O$, 0.5 $ZnSO_4.7H_2O$, 0.1 $MnSO_4.H_2O$, 0.05 $CuSO_4.5H_2O$, 0.04 $CoCl_2.6H_2O$, made up in 0.6 N HCl.
5 mL was added to CDYE (Section 2.2.2).

2.2.8 Top Agarose

Top agarose contained 1% (w/v) tryptone, 0.5% (w/v) NaCl, 0.8% (w/v) agarose. Before use the top agarose was supplemented with $MgSO_4.7H_2O$ to a final concentration of 10 mM.

2.2.9 Fungal Growth conditions

Fungal strains were grown at 22°C in potato dextrose broth or on potato dextrose agar plates (Section 2.2.5) until a suitable level of growth was attained. Cultures were then maintained at 4°C.

Endophyte strains were purified by serial plating on PD agar. *P. paxilli* was purified by plating a spore suspension for single colonies.

Spore suspensions of *P. paxilli* were made by growing cultures on ACM plates for three days and resuspending a block of mycelia in 0.01% (v/v) Triton X-100. The spore concentration was determined using a haemocytometer slide.

2.2.10 Bacterial Growth Conditions

E. coli cultures were grown at 37°C in LB broth or on LB agar plates (Section 2.2.3). Where necessary media was supplemented with ampicillin to a concentration of 100 μg/mL. Cultures were stored at -70°C in 50% (v/v) glycerol.

2.3 DNA isolation

2.3.1 Plasmid DNA

Plasmid DNA was isolated and purified using either a BioRad Quantum Prep® plasmid miniprep kit or a Qiagen plasmid mini kit according to the manufacturer's instructions.

2.3.2 Genomic DNA

Genomic DNA was isolated from freeze-dried mycelia or protoplasts using the methods of Möller, et al., (1992), Byrd, et al., (1990) or Yoder, (1988). Genomic DNA was isolated using both large scale and mini-prep scale.

DNA was extracted from fungi and endophyte-infected perennial ryegrass using the method of Möller et al (1992). Freeze dried material used for DNA extraction was ground to a fine powder in liquid nitrogen and resuspended in 10 mL of TES buffer (100 mM Tris, 10 mM Na_2EDTA, 2% SDS, pH 8.0). Proteinase K (2 mg; Roche) was added and the mixture incubated at 60°C for 1 h. NaCl and CTAB (10%) were added to a final concentration of 1.4 M and 1%, respectively and the solution incubated at 65°C for 10 min. Chloroform (14 mL) was added, the solution mixed thoroughly and incubated on ice for 30 min then centrifuged at 16,060 g for 10 min. Ammonium acetate (4.5 mL of 5 M) was added to the aqueous phase, which was mixed gently and placed on ice for 30 min, then centrifuged as above. The supernatant was mixed gently with 10.1 ml of isopropanol, and the DNA precipitated on ice for 15-30 min. The sample was centrifuged for 10 min as above, and the DNA pellet was washed with 70% ethanol, dried and resuspended in 500 μL of H_2O.

Genomic DNA was extracted from fungal isolates using the method of Byrd et al (1990). Freeze-dried mycelium (100 mg) was ground to a fine powder under liquid nitrogen and suspended thoroughly in 10 mL of extraction buffer (150 mM Na_2EDTA, 50 mM Tris, 1% sodium lauroyl sarcosine, 2 mg/mL proteinase K, pH 8.0). The solution was centrifuged at 2,000 g for 10 min at 4°C. The supernatant was incubated at 37°C for 20 min, extracted with an equal volume of Tris-equilibrated phenol (Amersham or Invitrogen) and spun at 20,200 g for 15 min at 4°C. The aqueous phase was extracted in a similar manner with equal

volumes of phenol/chloroform and finally with chloroform. The aqueous phase was centrifuged at 25,000 g for 20 min and the DNA was precipitated with one volume of isopropanol at 4°C for 10 min. The DNA was spun by centrifugation at 8,000 g for 10 min at 4°C and the resulting DNA pellet was washed with 70% ethanol, dried and resuspended in 500 μL of H_2O.

To isolate genomic DNA from fungal protoplasts the protoplasts were generated as in Section 2.9 by digestion of mycelia overnight in Glucanex (Chemcolour Industry). The protoplasts were filtered through a nappy liner (Chux), an equal volume of STC buffer (1 M sorbitol, 50 mM $CaCl_2$, 50 mM Tris-HCl, pH 8.0) added and the protoplasts pelleted by centrifugation at 8,000 g for 5 min. The protoplasts were resuspended in 0.5 mL of STC buffer, then 5 mL of extraction buffer that contained 50 μL RNase (1 mg/mL) was added. The method of Byrd, et al., (1990) was followed as above with the DNA precipitated with isopropanol directly after the chloroform extraction.

DNA was isolated on a small scale using modifications of the Yoder (1988) method as described. Freeze-dried mycelium (10 mg) was ground to a fine powder with liquid nitrogen in a microcentrifuge tube and homogenised with 500 μL of buffer (100 mM LiCl, 10 mM Na_2EDTA, 0.5% SDS, 10 mM Tris-HCl, pH 7.4). This mixture was extracted with Tris-equilibrated phenol (500 μL; Amersham or Invitrogen) and centrifuged for 15 min. The aqueous phase was extracted with an equal volume of chloroform, and the DNA precipitated with 800 μL of 95% ethanol. The precipitated DNA was washed with 70% ethanol, dried and resuspended in 50 μL of MilliQ H_2O.

2.3.3 Lambda DNA

Lambda DNA was isolated from a plate lysate using a PEG precipitation and phenol/chloroform purification. Phage were plated for confluent lysis on LB agarose plates (Section 2.2.3), overlaid with 5 mL of SM buffer (100 mM NaCl, 8 mM $MgSO_4.7H_2O$, 0.01 % (w/v) gelatin, 50 mM Tris-HCl, pH 7.5) and left overnight at 4°C. The lysate was collected and treated with DNase and RNase (final concentrations of 1 μg/mL) and incubated at 37°C for 30 min. A 5 mL aliquot of PEG solution (20% w/v PEG 6000, 2 M NaCl) was added, the solution mixed and incubated on ice for 1 hour. The phage was pelleted by

centrifugation at 5,800 g for 30 min at 4°C. The pellet was resuspended in 0.5 mL of SM buffer with 5 μL of 10% w/v SDS and 10 μL of 250 mM Na_2EDTA (pH 8.0) and incubated at 68°C for 15 min. The phage coat and associated proteins were removed by extracting with an equal volume of Tris-equilibrated phenol (Amersham or Invitrogen). The aqueous phase was further extracted with equal volumes of phenol and chloroform and finally with an equal volume of chloroform. The DNA was precipitated by the addition of an equal volume of isopropanol, incubated on ice for 30 min and spun for 10 min at top speed in a microcentrifuge. The DNA pellet was washed with 70% ethanol and resuspended in 50 μL of H_2O.

2.4 DNA manipulation

2.4.1 DNA quantification

DNA was quantified using a comparison to the Low Mass Ladder (Invitrogen) or by fluorometric quantification on a DyNA Quant (Hoefer) according to the manufacturer's instructions.

2.4.2 Restriction endonuclease digestion of DNA

Restriction endonuclease (RE) digests were carried out in the commercial buffer recommended by the manufacturer. Digests typically contained an excess of enzyme (~3-10 units of restriction enzyme/μg of DNA) and were performed in a water bath at the recommended temperature. A small aliquot of digested DNA was checked on a mini gel to ensure complete digestion.

Endophyte genomic DNA was typically digested overnight in a 200 μL volume containing 1-2 μg of genomic DNA and 1 x BSA (New England Biolab). A sample (10 μL) was checked for complete digestion by separating out on a minigel before the DNA was concentrated by ethanol precipitation. Typically genomic DNA required for Southern analysis was resuspended in 30 μL of H_2O and 20 μL of SDS loading dye (1% SDS, 0.02% bromophenol blue, 20 % sucrose and 5 mM Na_2EDTA) was added for loading.

DNA embedded in agarose plugs (Section 2.4.5.2) was first equilibrated three times in 10 mM Tris-HCl (pH 8.0). The plugs were then equilibrated for 30 min at 37°C in reaction buffer containing 1 x appropriate commercial buffer and 1 x BSA (New England Biolab). The reaction buffer was replaced with reaction buffer containing 100 - 200 units of RE equilibrated on ice for 10 min and then allowed to digest overnight at the recommended temperature.

2.4.3 DNA purification and precipitation

DNA was purified by phenol/ chloroform extraction to inactivate enzymes and remove impurities that may interfere with further manipulations. The DNA solution was extracted with an equal volume of Tris-equilibrated phenol (Amersham or Invitrogen) and chloroform, vortexed and spun for 5 min in a microcentrifuge. The aqueous phase was extracted with 1-2 volumes of chloroform as above. The DNA was then precipitated with ethanol or isopropanol.

PCR products that required sequencing or those used for radioactive hybridisation, were purified using the Qiagen minElute® PCR purification system according to the manufacturer's instructions.

Fragments that required extraction from an agarose gel were excised under long wavelength UV light and purified using the Qiagen gel extraction purification kit according to the manufacturer's instructions.

DNA that required concentrating was precipitated using 1/10 volume of 3 M sodium acetate (pH 7.0) and either 2.5 volumes of 95% ethanol or 0.6 volumes of isopropanol. The solution was mixed by inversion and incubated on ice for a minimum of 15 min. The DNA was pelleted in a microcentrifuge for at least 10 min. The DNA pellet was washed with 70% ethanol and dried at 37°C before resuspending in either H_2O or TE (10 mM Tris-HCl, 0.1 mM Na_2EDTA, pH 8.0).

2.4.4 Subcloning

PCR products that required subcloning were ligated into the pGEM®-T or pGEM®-T easy vectors (Promega) according to the manufacturer's instructions. Other vectors were prepared for subcloning by restriction enzyme digestion of the vector DNA (typically 2-5 μg) followed by incubation at 37°C with 0.1 unit of calf alkaline phosphatase (Roche) for 30 min. Na_2EDTA (final concentration of 5 mM), SDS (final concentration of 0.5%) and proteinase K (final concentration of 50 μg/mL, Roche) were added and the mixture was incubated at 56°C for 30 min. The DNA was purified by phenol/chloroform extraction and ethanol precipitated (Section 2.4.3). The vector was separated on an agarose gel and further purified by agarose gel purification (Section 2.4.3).

Ligations were performed in a 20 μL reaction volume as follows; 2 μL of ligation buffer (New England Biolab), 20 ng of vector, 0.1 -1 μL of T4 DNA ligase (New England Biolab) and a 2-3 fold molar excess of insert:vector. The ligation mixture was incubated overnight at 4°C.

E. coli strain XL-1 blue was transformed with ligation mixtures by electroporation. *E. coli* cells for electroporation were prepared by inoculation (1/100) of an overnight XL-1 culture into 1 L of LB (Section 2.2.3) and grown until mid-log phase (A_{600} 0.5-1.0). The cells were chilled on ice for 20 min and harvested by centrifugation at 4,000 g for 10 min (4°C). The cells were resuspended twice in ice-cold water (once in 1 L then in 500 mL, spinning as above) followed by two resuspensions in ice-cold 10% glycerol (20 mL then 4 mL). The cells were stored at -70°C.

Electrocompetent *E. coli* cells (40 μL) were mixed with 2 μL of ligation mix and incubated on ice for 1 min. The cells were placed in the base of an ice-cold 0.2 cm cuvette (BioRad), then pulsed at the following settings in a BioRad gene pulser; 25 μF, 2.5 kV, and 200 ohms. The cells were immediately resuspended in 1 mL of SOC (Section 2.2.6) and incubated at 37°C for 30 min - 1 h. The cells were plated at suitable dilutions onto selective LB plates.

2.4.5 Agarose gel electrophoresis

2.4.5.1 Standard electrophoresis

Agarose (Molecular biology grade; Roche or Invitrogen), at concentrations best suited for the desired separation, was melted in TBE (89 mM Tris, 89 mM boric acid and 2.5 mM Na_2EDTA, pH 8.2). Mini or overnight horizontal agarose gels were used to separate the DNA fragments by running in TBE buffer at 100 volts or 30 volts respectively. DNA samples were loaded in the wells with 1/5th volume of SDS dye mix (1% SDS, 0.02% bromophenol blue, 20% sucrose and 5 mM Na_2EDTA). Gels were stained in an ethidium bromide solution (1 μg/mL) for 10-15 min. The bands were visualised on an UV transilluminator and photographed with either an Alpha Innotech gel documentation system or BioRad gel documentation system. The DNA fragment sizes or concentrations were determined by comparison to known standards such as λ *Hin*dIII digested, 1 kb+ ladder, and the Low Mass Ladder (Invitrogen).

2.4.5.2 Pulse field gel electrophoresis

Chromosomal DNA and large DNA fragments were separated using the BioRad CHEF-DR® II system. Protoplasts were prepared as in Section 2.9. However, at the final step they were resuspended to a final concentration of 1 x 10^9 protoplasts/mL, mixed with an equal volume of 1.4 % low melting point agarose in GMB (0.9 M sorbitol, 125 mM Na_2EDTA, pH 7.5) and then allowed to set in plug moulds (BioRad) for 10-15 min. The plugs were incubated at 50°C for 18 h in SE buffer (2% SDS, 250 mM Na_2EDTA, pH 8.0), followed by 24 h incubation at 50°C in 10 mL of 10 x ET buffer (10 mM Tris, 500 mM Na_2EDTA, pH 8.0) with 20 mg of proteinase K (Roche) and 100 mg of sodium lauroylsarcosine (Sigma). The plugs were finally washed four times with 1 x ET buffer and stored at 4°C.

Plugs containing chromosomal DNA preparations were initially checked for quality and concentration on a 0.6 % chromosomal grade agarose (BioRad) gel in 0.5 x TBE using switch times of 100-1000 s, 100 volts for 19 hours at 10°C. Chromosomes were subsequently separated by one of two programmes. A mid-range separation used 0.7 % chromosomal grade agarose in 0.5 x TBE at 60 volts with the following switch times; 120 s for 24 h, 450 s for 20 h, 1500s for

72 h, 2100 s for 60.5 h all at 14°C. To separate the larger chromosomal DNA, the programme of Kuldau et al., (1999) was used as follows; 0.6% chromosomal grade agarose (BioRad) in 0.5x TBE at 40 volts at 10°C with switch times ramped from 1200-6000 s for 240 h. Digests of DNA embedded in plugs were performed as mentioned in Section 2.4.2. The DNA was separated on 1% molecular biology grade agarose in 0.5 x TBE at 200 volts with the following ramped switch times, 2-10 s for 18.4 hours then 10-16 s for 4.6 hours at 14°C. The gels were stained in an ethidium bromide bath and photographed as mentioned above.

2.4.6 Southern blotting

DNA was transferred to nylon membranes by the method based on that of Southern, (1975). DNA to be transferred onto membranes was separated overnight by agarose gel electrophoresis. The gel was gently agitated in the following series of solutions: blotting solution 1 (0.25 M HCl) for 15 min, blotting solution 2 (0.5 M NaOH, 0.5 M NaCl) for 30 min, blotting solution 3 (2.0 M NaCl, 0.5 M Tris, pH 7.4) for 30 min. The gel was washed for 2 min in 2x SSC (0.3 M NaCl, 0.03 M trisodium citrate) and assembled onto a blotting stand as follows: Two sheets of 3MM paper (Whatman) were used as wicks soaked in 20 x SSC (3 M NaCl, 0.3 M trisodium citrate), the sheets were covered with a film of plastic with a hole cut slightly smaller than the gel, the treated gel was placed on the stand followed by a piece of positively charged nylon membrane (Roche), two 3MM sheets (slightly smaller than the gel) soaked in 2 x SSC, 2 dry 3MM sheets then a stack of towels with a weight on top to keep the pile flat. The DNA was allowed to transfer to the membrane overnight. The membrane was washed briefly in 2 x SSC, then the DNA was fixed by UV irradiation of 120,000 μJoules/cm^2 using a Cex-800 UV-crosslinker (Ultra-Lum Inc.).

2.4.7 Radioactive hybridisation

DNA fragments used for hybridisation probes were amplified by PCR using primer combinations stated in Table 2.2 and then purified using the Qiagen MinElute® PCR purification kit. DNA (30 ng) was radioactively labelled using High Prime kit (Roche) with [α-^{32}P]dCTP (3,000 Ci/mmole, Amersham)

Table 2.2 Primer combinations for amplification of radioactive hybridisation probes

Amplification region/gene	primer 1 (5')	primer 2 (3')	Size bp genomic (cDNA)	Figure or application[3]
301	lol253	lol254	567	3.43
Actin (grass[1])	actinF	actinR	~790	3.13
AT-Rich	lol186	lol226	785	3.37
CY28 (*ltmG*)	ggpps27	ggpps28	209	3.4, library screen
CY29 (*ggs1*)	ggpps27	ggpps29	272	3.4, library screen
ggs1	CYLp19-12	CYLp19-1	463	3.6, 3.8, 3.57
hph	pUChph3	pUChph4	545	3.18
ltmB	lol345	lol346	250	3.51, 3.52, 3.55
ltmC	lol190	lol193	644	3.53
ltmC	lol216	lol236	1237	3.55
ltmC	lol189	lol190	360	library screen, 3.40
ltmE	lol356	lol341	677	3.53, 3.54
ltmE-J	lol292	lol265	2205	3.53
ltmG	lol3	lol1	407	library screen, 3.35, 3.53, 3.57
ltmG	lol79	lol27	1184	3.32
ltmG (cDNA)[4]	lol79	lol1	(524)	3.13, 3.52
ltmJ	lol205	lol206	242	library screen, 3.40, 3.51, 3.52, 3.53, 3.54, 3.55
ltmK	lol33	lol37	3358	library screen
ltmK	lol15	lol32	416	3.14
ltmK	lol29	lol63	1943	3.33, 3.36
ltmK (cDNA)[4]	lol32	lol8	(651)	3.13
ltmM	lol7	lol35	448	3.18, 3.51, 3.55
ltmM (cDNA)[4]	lol7	lol28	(780)	3.13, 3.27
ltmP	lol191	lol192	374	library screen, 3.40, 3.55
ltmP	lol196	lol198	426	3.40, 3.53, 3.54
paxM	mono3	mono4	~340	3.27
pks	lol103	lol109	1411	library screen, 3.35, 3.36
pks	lol128	lol129	519	3.57
Rua	lol4	lol23	383	library screen, 3.37, 3.57
Tahi	lol16	lol95	522	3.37, 3.57
tub2 (endo[2])	T1.1	T1.2	727	3.13

[1] actin (grass) refers to the amplification of a fragment from a ryegrass actin gene.

[2] tub2 (endo) refers to the amplification of a fragment from an endophyte β-tubulin gene.

[3] The figure that shows hybridisation with the probe, or the application it was used for.

[4] The cDNA products were cloned into pGEM-Teasy and then the fragment was excised from a gel.

Maps showing the fragment locations are in Fig. 3.5, 3.11, 3.26, 3.31, 3.34, 3.39, 3.41, 3.44, 3.50, and 3.54

Sizes in parentheses are RT-PCR products. These were cloned into pGEM-T easy (Promega) and the fragment was gel extracted. All other fragments are amplified from genomic DNA.

according to the manufacturer's instructions. After the labelling reaction a Probe Quant G50 Micro Column (Amersham) was used to remove the unincorporated isotope. The labelled probe was boiled to denature the DNA before addition to the pre-hybridised blot. Hybridisation was performed in either a container without shaking, or in a roller bottle in a Bachofer hybridisation oven at the stated time and temperatures. Membranes were washed and hybridisation signals detected by autoradiography on X-ray film (Fuji) for the required exposure time. X-ray film was developed using a 100Plus™ automatic X-ray processor (All-Pro Imaging Corp.).

2.4.7.1 Standard Southern hybridisation

Membranes were prehybridised for two hours at 68°C with 20-30 mL of hybridisation solution (50 mM HEPES, pH 7.0; 3x SSC; 18 mg phenol-extracted herring sperm DNA; 20 mg *E. coli* tRNA; 0.2% Ficoll-70; 0.2% bovine serum albumin; 0.2% polyvinylpyrrolidone; per L). The boiled radioactively labelled probe was added and allowed to hybridise overnight for 18 to 20 hours. Membranes were washed three to four times at 50°C in 2x SSC, 0.1% SDS, for a total of ~1-2 h.

2.4.7.2 Low-stringency Southern hybridisation

The membrane was prehybridised for 4 hours in 10x Denhardts (0.2% Ficoll; 0.2% polyvinylpyrrolidone; 0.2% bovine serum albumin; Sambrook et al, 1989) then hybridised for 48 h at 37°C with boiled radioactively labelled probe in 43% (v/v) formamide, 5x Denhardts, 5x SSC, 0.1% SDS, 10 μg/mL phenol-extracted herring sperm DNA and 50 mM sodium phosphate (Maniatis, et al., 1982). Post-hybridisation, the membrane was washed at room temperature in two changes of 2x SSC, 0.1 % SDS for 10 min, then three 15 min washes in 2x SSC, 0.1% SDS at 50°C.

2.4.7.3 Northern hybridisation

The *ltmG*, *ltmM* and *ltmK* fragments used for Northern analysis were amplified from a random primed cDNA pool generated from mRNA of endophyte-infected ryegrass. The fragments were cloned into the pGEM-T easy vector then excised from an agarose gel. The plant actin and endophyte *tub2* sequences were amplified from their respective genomic DNA. Northern blots were

prehybridised in 5x Denhardts, 0.5% SDS and 5x SSPE (one L contains: 43.8 g NaCl, 6.9 g $NaH_2PO_4.H_2O$, 1.85 g Na_2EDTA, pH 7.4), for two hours at 68°C then hybridised overnight with a boiled radioactively labelled fragment. The membranes were washed at room temperature for 20 min in 1x SSC, 0.1% SDS, then three 15 min washes at 68°C in 0.2x SSC, 0.1% SDS.

2.4.7.3 Stripping radioactive membranes

Membranes that were hybridised multiple times were stripped of their radioactive signal by 3 – 4 washes in boiling 0.1 % (w/v) SDS. After stripping, the membranes were exposed to X-ray film for 1 – 2 days to determine that the membrane was free of detectable signal.

2.5 Library screening

A library of *N. lolii* Lp19 genomic DNA in vector λGEM-12 was made as described in Dobson, (1997). The insert DNA was subjected to partial *Mbo*I digestion, size fractionated (9 – 23 kb), partially end filled and ligated into partially end filled *Xho*I sites within the lambda arms. This library has been used successfully for the isolation of clones containing following genes; *thi* (Zhang, 2004), *prt1* (McGill, 2000), *prt2* (McGill, 2000), *lpsA* (Panaccione, et al., 2001).

The library was titred using the *recA*⁻ *E. coli* host strain KW251 grown overnight in LB broth (Section 2.2.3) supplemented with 10 mM $MgSO_4.7H_2O$ and 0.2% (w/v) maltose.

Phage (100 μL) diluted in SM buffer (Section 2.3.3) to an appropriate concentration was mixed with 100 μL of KW251 and incubated at 37°C for 30 min. The phage/KW251 mixture was added to 3 mL of Top agarose (Section 2.2.8) equilibrated to 50°C and poured onto LB plates. The plates were incubated at 37°C for 6 to 8 hours until small plaques were visible. The plates were stored at 4°C wrapped with parafilm.

Filters (Hybond N+, Amersham), marked asymmetrically, were placed on the phage plates for 1 min and the marks transferred to the plates for later alignment. The filters were placed (DNA side up) on filter paper moistened with blotting solution 2 (Section 2.4.6) for 2 min, blotting solution 3 (Section 2.4.6) for 5 min and finally on 2x SSC (Section 2.4.6) for 2 min. The DNA was fixed by UV-crosslinking using a Cex-800 UV crosslinker (Ultra-Lum Inc) using 120,000 μJoules/cm^2. Hybridisation was performed as described above (Section 2.4). Positive plaques were picked, using a cut off P1000 blue tip, into 0.5 mL of SM buffer containing 50 μL of chloroform, where there was alignment of plaques on the plate to hybridising plaques detected on the X-ray film. Positive plaques were screened twice more using the procedure above.

The library was screened with the following probes (Table 2.2), CY28, CY29, *ltmG* (amplified with primers lol3 and lol1), *ltmK* (amplified with primers lol33 and lol37), Rua, *pks* (amplified with primers lol103 and lol109), *ltmC* (amplified with primers lol189 and lol190), *ltmP* (amplified with primers lol191 and lol192), *ltmJ* (amplified with primers lol205 and lol206). Each fragment was radioactively labeled as described in Section 2.4.7.

DNA isolated from positive clones using the method in Section 2.3 was digested with appropriate enzymes (usually at least one of *Bam*HI, *Eco*RI, *Hin*dIII, or *Sst*I) and the fragments separated by agarose gel electrophoresis. The resulting gel was Southern blotted (Section 2.4.6) and hybridised (Section 2.4.7) with the appropriate hybridisation probes. The banding and hybridisation pattern of each clone was compared to genomic hybridisation results.

DNA from isolated lambda clones was sequenced using either T7, Sp6 or gene specific primers. Fragments of interest were randomly cloned into a pUC118 base vector and sequenced (Section 2.4.4).

2.6 RNA isolation and analysis

2.6.1 Isolation of RNA

Total RNA from mycelia or ryegrass pseudostems was isolated using Trizol reagent (Invitrogen). Approximately 1 g of tissue was ground to a fine powder with liquid nitrogen, in a mortar with a pestle. The tissue was then mixed with Trizol reagent (10 mL) and centrifuged at 12,000 g for 10 min at 4°C. Chloroform (2 mL) was added to the supernatant, mixed thoroughly for 15 s, left at room temperature for 3 min, then centrifuged as above for 15 min. Isopropanol (5 mL) was added to the aqueous phase, incubated at room temperature for 10 min, then spun as above for 10 min. The RNA pellet was washed with 10 mL of 75% ethanol, centrifuged at 7,500 g for 5 min and allowed to air dry in the fumehood. The RNA pellet was resuspended in 200 μL of DEPC-treated H_2O. The purity and quantity of the RNA was determined using spectrophotometer readings at A_{260}/A_{280}.

2.6.2 cDNA analysis

Messenger RNA was isolated by oligo(dT) affinity chromotography using a GenElute™ mRNA miniprep kit (Sigma). Approximately 50-100 ng of mRNA was denatured at 65°C for 10 min in the presence of random primers (1.8 nmol; Roche) in a 12 μL volume. The reaction volume was increased to 20 μL by the addition of 4 μL of 5x reaction buffer (Roche), 10 mM DTT, 1 mM of each dNTP, 8 units of RNase inhibitor (Roche) and 50 units of Expand Reverse Transcriptase (Roche). The reaction was incubated at 30°C for 10 min, then at 42°C for 45 min. Gene-specific amplification of cDNA dilutions was carried out in a 25 μL reaction volume that contained 1x *Taq* polymerase buffer (Roche), 50 μM of each dNTP, 200 nM each primer and 0.5 units of *Taq* polymerase (Roche). Cycle conditions were typically one cycle of 94°C for 2 min, followed by 35 cycles of 94°C for 15 s, 60°C for 30 s and 72°C for 45 s (or 1 min/kb) with one cycle of 72°C for 10 min. The products of the PCR reactions were separated on a 2% agarose gel with TBE buffer.

Gene expression was also determined using SuperScript™ One-step RT-PCR with Platinum *Taq* (Invitrogen) according to the manufacturer's instructions with 1 μg and/or 0.1 μg of total RNA.

2.6.3 Northern blotting

Total RNA (up to 17 μL) was mixed with 35 μL loading buffer that contained 17.5 μL formamide, 6.5 μL formaldehyde (6.9%), 4μL 10x MOPS buffer (per L; 41.9 g MOPS free acid, 6.8 g of sodium acetate, 10 mM Na_2EDTA, pH 7.0), 1.6 μg ethidium bromide. Samples were denatured for 15 min at 65°C. RNA was separated at 100 volts on a formaldehyde agarose gel containing 1.2% agarose, 6.2% formaldehyde and 1x MOPS buffer. The RNA was transferred to positively charged nylon membrane (Roche) via capillary transfer with 20 x SSC overnight. Hybridisations were performed as in Section 2.4.7.

2.7 DNA sequencing and Bioinformatics

The DNA sequence data were generated by sequencing purified PCR fragments or fragments cloned into pUC-based vectors, pGEM-based vectors or lambda clones. DNA fragments were sequenced using Big-Dye (version 3) chemistry (Applied Biosystems) with sequence-specific oligonucleotide primers generated by Invitrogen or Sigma Genosys. The products were separated on an ABI Prism 377 or 3730 sequencer (Perkin-Elmer).

Sequence data were assembled into contigs using Sequencher version 4.1 (Gene Codes) and analysed using the Wisconsin Package version 9.1 (Genetics Computer Group). Sequence comparisons were performed through Internet Explorer version 6.0 at the National Center for Biotechnology Information (NCBI) site (http://www.ncbi.nlm.nih.gov/) using the Brookhaven (PDR), SWISSPROT and GenBank (CDS translation), PIR and PRF databases employing algorithms for both local (BLASTX and BLASTP) and global (FastA) alignments (Altschul, et al., 1990; Altschul, et al., 1997; Pearson, Lipman, 1988). Sequences were also compared to the fungal genome sequences based at the Broad Institute (http://www.broad.mit.edu/annotation/fungi/fgi). E-values from BLAST anaysis were generally considered significant at 1e-5 and below. Sequence annotation was performed with the MacVector software version 7.2. The generated sequence is shown in Sequencher and MacVector formats in Appendix 5.3.

Introns contained in genomic sequences were identified as regions of a gene that introduced frame-shifts, stop codons, or gaps in the alignment with a polypeptide sequence. The intron sequence was predicted based on the removal of the 5' (GT) to the 3' (AG) splice sites, which resulted in an open reading frame. The predicted introns were confirmed by cDNA analysis using primers within the coding region to amplify across the introns. The cDNA PCR products were purified and sequenced with gene-specific primers. The cDNA sequence was compared and aligned to the genomic sequence using the GCG programme GAP.

The upstream region from the ATG of each *ltm* gene and *ggs1* was screened through Safari version 1.2.4 at the site for Regulatory Sequence Analysis Tools at http://rsat.ulb.ac.be/rsat/ (van Helden, 2003; van Helden, et al., 2000a; van Helden, et al., 1998; van Helden, et al., 2000b). The sequences were subjected to the pattern discovery algorithms, oligo-analysis (oligonucleotide size selected from 5-8 bases), dyad-analysis (oligonucleotide size 3 with spacing from 0-20 bases) and consensus (matrix lengths from 7-10). Selected motifs were aligned to the promoter sequences using 'dna-pattern' an algorithm for pattern matching. The location of each motif was shown using 'feature map'.

Multiple sequence alignments were performed using the CLUSTAL-X (Thompson, et al., 1997) and CLUSTAL-W software (Thompson, et al., 1994).

The phylogenetic trees of the fungal GGPP synthase were drawn based on 205 conserved positions using SplitTrees (Huson, 1998) to build a split decomposition graph and neighbour joining graph

2.8 PCR analysis

Oligonucleotide primers cited throughout this thesis are listed in Table 2.3. Primers were synthesised by Invitrogen or Sigma Genosys and resuspended to a stock concentration of 100 pmol/μL. Primers used for PCR stocks were diluted

Table 2.3 Primers cited within this thesis

Primer name	Sequence 5'-3'	Used for
ActinF	GCTGTTTTCCCTAGCATTGTTGG	probe
ActinR	ATAAGAGAATCCGTGAGATCCCG	probe
CY4	GCTTGGATCCGATATTGAAGGAGC	amplify *hph*
CY5	TTGGATCCGGTTCCCGGTCGGCAT	amplify *hph*
CY9	AGCGGATCCCTTCTAAGCGTTGGTGGATTG	*paxM* promoter
CY11	AGCGGATCCATAACTTGAAACTCGGCCTT	*paxM* promoter
CY16	ATICAYYTITGYYTIATG	Degenerate PCR
CY17	GARACIGCIAAYMGIGCITAYTA	Degenerate PCR
CY18	CKRAAIARISCICCIGTYTT	Degenerate PCR
CY19	TAIACRTTYTTRCARTCRTTYAG	Degenerate PCR
CY20	GAGCTCCAGGAGACACAGAC	E8 *ggs1* PCR
Endo1	ACCCTTTGACTACGTGG	real-time PCR
Endo2	AGATGTTGTGGGCGAC	real-time PCR
ggpps18	AGTATGGCTTACTCAAGCTG	*paxM* PCR
ggpps27	CAYMGIGGTCARGGTATGGA	Degenerate PCR
ggpps28	TTCATRTAGTCGTCICKTATYTG	Degenerate PCR
ggpps29	AACTTTCCYTCIGTSARGTCYTC	Degenerate PCR
CYLp19-1	CACCATTTCGAGGTAGTC	*ggs1* PCR/probe/RT-PCR
CYLp19-6	CATGAGGAATCCCACGAC	PCR
CYLp19-7	GAGGTACATGTGAAATGGAAC	*ggs1* PCR
CYLp19-8	ATCTCCCAATCCCATTCC	*ggs1* PCR
CYLp19-12	GGATTTGAATTACACGCC	*ggs1* probe
CYLp19-13	AAAGATGTTGTGGGCGAC	seq
CYLp19-15	ACGTGGGATCCTATTGTCGCTGGGC	*ggs1* PCR
CYLp19-16	CGTCGCCCACAACATCTTTG	*ggs1* IPCR/PCR/RT-PCR
CYLp19-17	CTACATCTACTTTGTCGCCC	seq
CYLp19-18	AGCGAAAGTCCTTGCCCGGA	*ggs1* IPCR/PCR/RT-PCR
CYLp19-19	TTGTCGTCGGTCCATGAGCG	seq
CYLp19-20	CGCTCATGGACCGACGACAA	*ggs1* PCR
CYLp19-22	TCACAAGATGGCGTCGTCG	*ggs1* RT-PCR/PCR
CYLp19-23	TTTTGAAACAACCATTCTGTCC	PCR
lol1	TGGATCATTCGCAGATAC	*ltmG* probe/RT-PCR
lol2	GTGTGAGATTAAGACGTC	PCR
lol3	ACCGACGCCATTAATGAG	*ltmG* probe/IPCR/RT-PCR
lol4	ATAGTCTAACTAGAGGGC	Rua probe
lol7	ACTGGGCATCTTCCATAG	*ltmM* probe/RT-PCR/PCR
lol8	CTTGAAGTGTGATACCGG	*ltmK* probe
lol9	AACGACGCAATGATTCGC	IPCR
lol13	ACGCCATATCTATTGCCG	IPCR
lol14	ATTAGAGGCACCGAACGC	*ltmM* RT-PCR
lol15	ATCAAGCTGGCTATCCTC	*ltmK* probe/RT-PCR/PCR
lol16	AGTCTTTCCTAGCGTAGG	Tahi probe
lol17	AAATAATGGGCAAGGAGC	KO PCR
lol18	TGGGATTTTGGAAATGGC	KO PCR/IPCR
lol20	CAAAGAATTCGCTTGAGCACCCGA	*ltmM* PCR
lol23	TAGATAGGGAAGTTATGC	Rua probe
lol27	GACGTCTTAATCTCACAC	*ltmG* probe
lol28	CGACCTTGCCCATTATTT	*ltmM* probe/RT-PCR
lol29	GTCTTGATCGTCTGCATC	*!tmK* probe/RT-PCR
lol32	TGTCCGTGCATCCATTGT	*ltmK* probe/RT-PCR
lol33	GGTTAGTCTTAGCTAGAG	*ltmK* probe
lol34	CATAGAGCTAGCTAGAGT	PCR
lol35	GTTCGGTGCCTCTAATAC	*ltmM* probe/RT-PCR/PCR
lol37	GATGTTAAGGCGTGCATG	*ltmK* probe
lol40	TGAGCCTAAGGAGATAGC	IPCR
lol43	GAGGATAGCCAGTTGAT	*ltmK* RT-PCR

Primer name	Sequence 5'-3'	Used for
lol48	GATTGGTACCTTGAAGTCGCTAGT	KO PCR
lol49	GTAGGGTACCTCTAGTACTGCCTCT	KO PCR
lol54	TATGAGGATCCTTGCTTGTCTGTTT	*ltmM* PCR
lol63	TAGCGAATCATTGCCGTCG	*ltmK* probe/RT-PCR
lol79	ATGGCTGCCAATGACTTTCC	*ltmG* probe/RT-PCR
lol95	GTAAGCGGTTAAAAGGGA	Tahi probe
lol103	AAGAGGGCGCCCACCAATTC	*pks* probe
lol109	AGACCCTGGTCCAAGTGG	*pks* probe
lol110	CACTTCGCATATTCTTCAGG	*pks* RT-PCR
lol111	TGGTGTCATTGGAACACGG	*pks* RT-PCR
lol128	GCAAGTAGTGGTAGGCATCG	*pks* probe
lol129	TCTGCATCACTCCACAGGGT	*pks* probe
lol135	AGGCCATTTTCGACAGTTGT	PCR
lol147	CCAGCAAGCATGCACATTAC	PCR
lol148	TGCGTGAGAGATAAAGCAAG	PCR
lol156	GCACAAACAATAAATTCGGCCAA	*ltmG* PCR
lol157	AATTTGCCCTCTGTTAAATCCTC	*ltmG* PCR
lol158	GTGATCGGTGCTGACGGGGTCCA	*ltmM* PCR
lol159	TATCGCCATATTTGCTCCTTGCCC	*ltmM* PCR
lol160	ATATTGAATTGCTGCGTGAGGAG	*ltmK* PCR
lol161	AGAGGCCAAGAAGCGGCCTGGACA	*ltmK* PCR
lol186	AAAGCTAAGGGCTATAAGAG	AT-Rich probe
lol189	AAAGACATGTCTTGTCCCTC	*ltmC* PCR/probe
lol190	GATTCCTGACTGTCCATACTC	*ltmC* PCR/probe
lol191	CCAAGGAGGTTTTGAATGTA	*ltmP* PCR/probe
lol192	TTGGATGAGCTCAATCATGC	*ltmP* PCR/probe/RT-PCR
lol193	TGAGTATATGTTCTTGCAATCA	*ltmC* probe/RT-PCR
lol194	GAACTCGTAGCGCAGGAGCA	*ltmJ* PCR
lol195	TTCTCTTCGGAGGCTCTCCTT	*ltmP* probe
lol196	TGGACATGGATCTGATTGTC	*ltmP* probe
lol197	TGGATTGACAGAAATCCAAG	*ltmP* RT-PCR
lol198	TGTAGCACGGGTAGCTAGAT	*ltmP* probe
lol199	TTGCGCATCGTACGCTAGGA	IPCR
lol202	GGATGAAGAAAATCCACGAG	IPCR
lol203	AGACGATCTGTTAGGCCGAT	IPCR
lol205	CCAAGCATCGATTTGTCACC	*ltmJ* PCR/probe
lol206	AATCTGATCGCCATCTTTGC	*ltmJ* PCR/probe
lol207	TCGCTTATCAGGAAGTTCAA	*ltmB RT-PCR*
lol209	GAATAGCTCAAGACTCAGAA	IPCR
lol210	AAGCTGGCTGTTAAAGGGTC	IPCR
lol211	TATTAGGGAGCGAACTTCAC	IPCR
lol213	AAGAGGGCCGCAATTTCGAT	IPCR
lol214	GGTGACAAATCGATGCTTGG	*ltmJ* RT-PCR
lol216	AGATGACATCTGGAGCATGG	*ltmC* probe/RT-PCR
lol218	GCTCCTGAAGCAATAGGTGC	*ltmD* RT-PCR
lol220	TCGTTTCCTTTCAACATACG	*ltmP* RT-PCR
lol221	GAGAATGTTAATGTTGCACGC	*ltmP* RT-PCR
lol222	GCGTGCAACATTAACATTCTC	IPCR
lol223	TCGTTGCTGTTAGGTCGCAG	*ltmQ* RT-PCR
lol226	GCTTAAAAGAGCTCTCTATAGC	AT-Rich probe
lol227	CAAGACTCCAGAGGATAAGG	*ltmB RT-PCR*
lol228	CCTCAAGCAACGAGATTTGAC	*ltmD* RT-PCR
lol234	GTTAAAGTCGTCAAGCTAGG	*ltmQ* RT-PCR
lol235	ATTCCACCATGGCATCTGGAGCATGGCTCG	*ltmC* complementation
lol236	CTTAAGCGAATTCTACCTTGTGGGTC	*ltmC* probe/complementation
lol246	ATGATCTGGCACTCTAGGAG	*ltmB RT-PCR*
lol253	CCTTACTTATGCCTAACCCT	301 probe
lol254	GTAAGTATACCTAGGCTACC	301 probe
lol263	ATTCAAATGGCTACTCCCAG	*ltmQ* RT-PCR

Primer name	Sequence 5'-3'	Used for
lol265	CGAGGAAGTAGTTTCAGGGC	*ltmE-J* probe/*ltmJ* RT-PCR
lol266	GACAGCGTCTTGAGGGAATC	*ltmJ* RT-PCR
lol274	GGCACCTATTGCTTCAGGAGC	*ltmD* RT-PCR
lol275	CTCACATTACACCATGATTGC	*ltmD* RT-PCR
lol277	GCTTTCACAGATTCTTCTCC	*ltmQ* RT-PCR
lol278	GAAACTGCCAATCGAGCATA	*ltmC* PCR/RT-PCR
lol279	TTCTTGCAATCATTTTGCAATTG	*ltmC* PCR/RT-PCR
lol280	ATGGCTGTCATTCATACAACAGCTATG	*ltmP* PCR/RT-PCR
lol281	AGCGTCCCGGACAGGCATATCTCCCA	*ltmP* PCR/RT-PCR
lol282	CAGAGGTTTAACCCTCTTGACGC	*ltmQ* PCR/RT-PCR
lol284	TGAAACCATGGCTAGCGACTTCAAGGT	*ltmM* complementation
lol285	ACATTGAGCTCCTAGTTGGAAAATGGG	*ltmM* complementation
lol289	GTCGCCGATGTATCGTATTG	*chsV* PCR/RT-PCR
lol291	AGATCCTCCATGGAGGAACG	*ltmE* RT-PCR
lol292	GCGTCAAGTCTACATACCTG	*ltmE-J* probe
lol293	ATATCTGGAACCATGCAGTG	*ltmJ* RT-PCR
lol310	CTTTGATGAAACTATCCACC	*ltmQ* RT-PCR
lol311	CTGGGAGTAGCCATTTGAAT	*ltmQ* RT-PCR
lol312	CTGGCCAGTCGTTTCCACGT	*chsV* PCR/RT-PCR
lol313	CTACCAGGACAGGCGTGACGTCC	*ltmQ* PCR
lol341	TTCCGCTTCCGAGTAGACTC	*ltmE* PCR/RT-PCR/probe
lol343	ATCGATGCCAAGAGTGATGC	*ltmE* RT-PCR
lol345	AACATCGCCTGGGAGCTCGTATA	*ltmB probe/PCR*
lol346	CGCAGGTCCTATTTCCATCGC	*ltmB probe/PCR*
lol356	CCGAGTTTGATGACCTGCTG	*ltmE* PCR/RT-PCR/probe
lol359	GAATTATGTTACTCTTGGGG	*ltmD* PCR/RT-PCR
lol360	AAGTTGGCACATAGGTCTTC	*ltmD* PCR
mono3	CTGCGCCGACAAGAAGATCC	*paxM* probe
mono4	AGCATATCGAACCGCTAAGC	*paxM* probe
paxMNotI	GCCATGCGGCCGCTTGTTGA	*paxM* PCR
pII99-1	CTTTGAACAGCGACGGTC	Seq
pII99-2	TTGAGTGAGCTGATACCG	Seq
pII99-3	GGCTGGCTTAACTATGCG	Seq
pII99-4	CCCAGAATGCACAGGTAC	Seq
prg3	GCTCACTGATTCCGTCCTTG	RT-PCR
prg4	CACTCATCTGCATTCTCAAC	RT-PCR
prg12	ACGTCTACTACAACGAGGCC	real-time PCR
prg13	CCTTGGCCCAGTTGTTACCG	real-time PCR
PUC forward	GTTTTCCCAGTCACGAC	Seq/PCR
pUChph3	CTGCATCATCGAAATTGC	*hph* probe
pUChph4	AAACCGAACTGCCCGCTGTTC	*hph* probe
PUC reverse	CAGGAAACAGCTATGAC	Seq/PCR
SP6	CCATTTAGGTGACACTATAG	Seq
T1.1	GAGAAAATGCGTGAGATTGT	*tub2* probe/RT-PCR
T1.2	CTGGTCAACCAGCTCAGCAC	*tub2* probe/RT-PCR
T7	TAATACGACTCACTATAGGG	Seq

to a final concentration of 10 pmol/μL and primers used for sequencing stocks were diluted to 3.2 pmol/μL. All primer stocks were stored at -20°C.

2.8.1 Standard PCR reaction conditions

The following standard PCR reaction conditions were used in experiments reported in this thesis unless otherwise stated: 5 ng genomic DNA, 1x *Taq* polymerase buffer (Roche), 50 μM each dNTP, 200 nM each primer, 0.5 U *Taq* polymerase in a reaction volume of 25 μL. Reactions were also scaled up to 50 μL. The PCR cycling conditions for each experiment is stated within the figure legends.

2.8.2 Degenerate PCR

The following degenerate PCR conditions were used: 5 ng genomic DNA, 1x *Taq* polymerase buffer (Roche), 50 μM each dNTP, 800 nM each primer, 0.5 U *Taq* polymerase in a reaction volume of 25 μL. The amplification conditions were 95°C for 2 min followed by 30 cycles of 95°C for 30 s, 45°C for 30 s and 72°C for 1 min, then 1 cycle of 72°C for 5 min. The annealing temperature was also increased to 47°C with a similar amplification result.

2.8.3 Inverse PCR (IPCR)

One microgram of genomic DNA was digested with the appropriate restriction enzyme as determined by Southern analysis. After complete digestion the DNA was purified using phenol extraction. A self-ligation was performed to a final concentration of 10 ng/μL of digested genomic DNA with 1x T4 DNA ligase buffer (New England Biolabs) and 40 units of T4 DNA ligase (New England Biolabs) in a 20-100 μL reaction. The self-ligated genomic DNA (50 ng) was used in a standard PCR reaction with primers designed in the opposite direction from each other. IPCR products were sequenced with sequence-specific primers or cloned into pGEM-T easy and sequenced. The sequence generated from cloned fragments was confirmed by comparison to sequence generated from a genomic library.

2.8.4 Colony PCR

Colonies were picked into a 25 μL standard PCR reaction mixture, then amplified with the following cycle conditions: one cycle of 94°C for 5 min;

followed by 30 cycles of 94°C for 15 s, 60°C for 30 s and 72°C for 1 min/kb; then one cycle of 72°C for 10 min.

2.8.5 Real-time PCR analysis

Endophyte-infected ryegrass genomic DNA was isolated by the method of Möller et al (1992). The DNA was resuspended to a concentration of 300 μg/μL boiled for 5 min, the concentration quantified using a Hoefer DyNAquant 200 fluorometer and diluted to 5 ng/μL. An endophyte-free ryegrass standard curve was prepared by diluting boiled G1138 DNA to a range of concentrations between 0.01-10 ng/μl. The endophyte standard curve was prepared by diluting boiled strain Fl1 genomic DNA to a range of concentrations between 0.015-1.5 ng/μL with G1138 genomic DNA at 5 ng/μL. Amplification of the endophyte and plant amplicons was performed on a Lightcycler (Roche) using the Lightcycler FastStart DNA Master^{PLUS} SYBR green I system (Roche) with primers endo1 and endo2 and prg12 and prg13 (Table 2.3) for the endophyte and ryegrass amplification respectively. Each reaction contained 2 μL of Roche Lightcycler FastStart DNA Master^{PLUS} SYBR green I system, 62.5 nM of each primer, and 2 μL of sample DNA in a final volume of 10 μL. The amplification programme was 95°C for 10 min, then 35 cycles of 95°C for 10 s, 60°C for 5 s and 72°C for 9 s or 10 s for ryegrass or endophyte amplification respectively using a single fluorescence acquisition. To obtain the melting curves the lightcycler was held at 95°C for 0 s, 65°C for 30 s, then ramped up to 95°C at a rate of 0.2°C/s using continuous fluorescence acquisition. Standard curves used to determine the endophyte and ryegrass DNA concentrations were generated from the genomic DNA standards using the above reaction conditions. The standard curve was plotted as log concentration of the DNA standards versus the cycle number of the amplified product. The standard curve was generated in duplicate and used externally. A single standard was run with each set of samples as a calibration for the standard curve. Each sample was analysed in duplicate for each primer combination. The quantification analysis was determined using the fit points selection with the noise bands set to 0.505. The melting curves were used to determine whether there was a specific amplification of each product and this was confirmed by running the PCR reactions on a 2% agarose gel. The endophyte DNA biomass was determined as

a percentage of ryegrass sample using the calculation (E/R)100 where E and R are the number of pg of endophyte and ryegrass DNA respectively in the sample.

2.8.6 Long template

Linear fragments used for transformation were amplified using the Expand long-template reaction mixture (Roche). The plasmid DNA was diluted to ~1 ng/μL and used in the following PCR reaction: 5 ng of plasmid DNA; 1x Expand Long Template buffer 1; 350 μM each dNTP; 150 nM each primer; 0.75 μL of Expand Long Template enzyme mixture (Roche) in a 50 μL reaction mixture. The amplification conditions were as follows: one cycle of 93°C for 2 min, followed by 30 cycles of 93°C for 10 s, 55°C for 30 s and 68°C for 6 min with a 20 s increment/cycle after cycle 12; then one cycle of 68°C for 10 min.

2.8.7 High Fidelity enzymes

PCR products that required amplification with a high-fidelity enzyme were amplified with either Expand High Fidelity (Roche) or with Platinum *Pfx* (Invitrogen) according to the manufacturer's instructions.

2.9 Fungal Transformations

2.9.1 Fungal protoplasts

Protoplasts were prepared using modifications of the methods described by Murray, et al., (1992), McMillan, et al., (2003), Yelton, et al., (1984), and Itoh, et al., (1994). Fungal cultures inoculated with ground mycelia (*N. lolii* and *E. festucae*) or 5 x 10^6 spores (*P. paxilli*) were grown in 25 mL of PD broth at 22°C until appropriate growth was attained. *P. paxilli* was typically grown for 2 days, *E. festucae* was grown for 6-10 days and *N. lolii* was grown for 10-14 days. The mycelia were washed three times in sterile water followed by a wash in OM buffer (1.2 M $MgSO_4.7H_2O$, 10 mM Na_2HPO_4, adjusted to pH 5.8 with 100 mM $NaH_2PO_4.2H_2O$). The washed mycelia were mixed with 15 mL of filter-sterilised (with a 0.45 μ arcodisc) OM buffer containing 10 mg/mL of glucanex (Chemcolour industry) and incubated overnight at 30°C with gentle shaking. The digested hyphae were filtered through a nappy liner (Chux) and the filtrate containing protoplasts overlaid with 2 mL of ST buffer (0.6 M sorbitol, 100 mM

Tris-HCl, pH 8.0). The protoplasts were banded at the interface by centrifugation at 3,000 g for 5 min. The protoplasts were removed from the interface and washed three times with 5 mL of STC buffer (1 M sorbitol, 50 mM $CaCl_2$, 50 mM Tris-HCl, pH 8.0) by centrifugation at 7,700 g. Protoplasts were resuspended in STC buffer to give a final concentration of 1.25 x 10^8 protoplasts per mL. Protoplasts (80 μL) were frozen at -80°C for long-term storage by adding 20 μL of PEG buffer (40 % PEG 4000, 50 mM $CaCl_2$, 1 M sorbitol, 50 mM Tris-HCl, pH 8.0).

2.9.2 Fungal transformation

Protoplasts of *E. festucae*, *N. lolii* or *P. paxilli* were transformed by the method of Vollmer and Yanofsky (1986), as modified by Itoh et al (1994). The deletion of the *ltmM* gene and subsequent complementation into the CYFII-M28 and CYFII-M61 background used 5 μg of appropriate DNA. The DNA used to delete *ltmM* was linearised by amplification with the Expand Long Template PCR System (Roche) (Section 2.8.6), while all other templates used were circular. DNA amounts for the complementation of the *P. paxilli* strains were 5 μg for plasmids that contained an integrated geneticin selectable marker or a co-transformation was used with 5 μg of pII99 and 15 μg of the target plasmid.

DNA was added to 80 μL of protoplasts with 20 μL of PEG buffer, 5 μL of heparin (5 mg/mL in STC buffer), 2 μL of spermidine (50 mM). After incubation on ice for 30 min, 900 μL of PEG buffer was added and the protoplasts were incubated at room temperature for 15 min. A 100 μL aliquot of the transformation mix was plated onto RG plates (2.4% potato dextrose broth, 0.8 M sucrose, 1.5% agar, pH 8.0) with a 3 μL overlay of RG-top media (2.4% potato dextrose broth, 0.8 M sucrose, 0.8% agar, pH 8.0). The plates were incubated at 22°C overnight and then overlaid with either hygromycin (Roche) or geneticin (Invitrogen) at the concentrations determined in Section 3.6. The transformants were allowed to grow at 22°C for 7 days (*P. paxilli*), 1 month (*E. festucae*) or 6 weeks (*N. lolii*).

2.9.3 Nuclear purification of transformants

Geneticin-resistant colonies of *P. paxilli* transformants were plated for single colonies onto PD media and grown at 22°C for 5 days. Single colonies were

picked into 75 μL of 0.01% Triton X-100, 50 μL was spread onto ACM media (Section 2.2.1) using glass-plating beads and grown at 22°C for 4 days. A spore suspension was created by resuspending a 4 cm agar block from the sporulating plates in 2 mL of 0.01% Triton X-100.

E. festucae and *N. lolii* transformants were nuclear purified by sub-culturing mycelia from the edge of a colony to PD media, containing either hygromycin (150 μg/mL) or geneticin (200 μg/mL) three consecutive times.

2.10 Plant growth, inoculations and analyses

2.10.1 Seedling inoculations

Seedling inoculations were carried out by the method of Latch and Christensen (1985). Endophyte-free seeds of perennial ryegrass 'Nui' cultivar were surface-sterilised by first soaking in 50% sulfuric acid for 30 min, rinsing three times with water, then 20 min in 50% commercial bleach, rinsed three times in sterile water then allowed to air dry. The seeds were germinated in the dark at 22°C for a week on 4% water agar. A small piece of a freshly grown fungal colony was inserted into a slit in the region between the mesocotyl and coleoptile of each seedling. Multiple seedlings were inoculated per culture. The inoculated seedlings were incubated at 22°C for 5-7 days in the dark followed by 5-7 days in the light. The plants were transferred into root trainers containing commercial potting mix and maintained in the greenhouse for 6-8 weeks before being tested for endophyte infection. Endophyte-infected plants were transferred into pots or bags containing potting mix and slow-release fertilizer (Osmocote) and watered regularly.

2.10.2 Immunoblotting to screen for endophyte infection

Plants were tested for endophyte infection using tissue-print immunoblotting, a modification of the method described by Gwinn, et al., (1991). Polyclonal antibodies were raised in New Zealand white rabbits that were injected with homogenised *N. lolii* mycelia (strain Lp5) (Christensen, et al., 1993). Grass tillers were cut transversely at the base and printed onto a nitrocellulose membrane for several seconds. The membrane was washed in blocking solution

(20 mM Tris, 50 mM NaCl, 0.5% non-fat milk powder, 10 mM HCl, pH 7.5) at room temperature for 2 h, transferred to primary antibody diluted 1:1000 in blocking solution and incubated overnight at 4°C with gentle shaking. The membrane was rinsed five times in blocking solution, then incubated for 2 h with shaking at room temperature in secondary antibody (goat anti-rabbit with an alkaline phophatase conjugate; Sigma) diluted 1:2000 in blocking solution. The membrane was developed in Fast Red chromogen (0.6% Fast Red in 200 mM Tris buffer mixed with 0.1% naphthol as-mx phosphate (Sigma) in 200 mM Tris buffer) for 15 min at room temperature with shaking and then rinsed with water to stop the reaction.

2.10.3 Aniline blue staining of endophyte

Plant epidermal strips were examined with a light microscope for the presence of endophyte by staining with aniline blue (0.05% aniline blue in lactic acid/glycerol/water 1:2:1). Epidermal strips from the outer most leaf sheath were briefly heated in the presence of the aniline blue stain and then examined using a compound microscope at 100x and 400x magnification. Typically hyphae were stained blue and could be observed between the plant cells.

2.11 Alkaloid Analysis

Endophyte-infected plant material (pseudostems) was freeze-dried and ground to a fine powder. A 50 mg sample was extracted for the alkaloid lolitrem B, ergovaline and peramine.

2.11.1 Lolitrem B analysis

Lolitrem B was analysed by a modification (Panaccione, et al., 2003) of the method of Gallagher, et al., (1984). Lolitrem B was extracted from 50 mg of endophyte-infected plant material with 1 mL of dichloroethane:methanol (9:1). The sample was agitated in a cell disrupter (FP 120 Savant FastPrep, BIO 101 Inc.) for 20 s, speed 5, mixed by rotation for 1 h, then centrifuged for 10 min at 3,000 g. An 8-20 μL sample was analysed by normal phase HPLC (Shimadzu LC-10A system) on silica columns (Altima 5 μ, 150 x 4.6 mm, Alltech Associates, Deerfield, II) with dichloromethane:acetonitrile:water (880:120:1)

running solvent at a flow rate of 1 mL/min at 28°C. Lolitrem B was detected by fluorescence (excitation 265 nm, emission 440 nm; RF-10A detector, Shimadzu). Lolitrem B elutes at ~5 min followed by smaller amounts of other lolitrems. The amount of lolitrem B was estimated by comparison of integrated peak areas with external standards of authentic lolitrem B (Class-LC10 software, Shimadzu).

2.11.2 Ergovaline and peramine extractions

Ergovaline and peramine were extracted by minor modifications of previously described methods (Panaccione, Schardl, 2003; Spiering, et al., 2002). Ergovaline was extracted from 50 mg of endophyte-infected plant material with 1 mL of isopropanol-lactic acid (50% (v/v) isopropanol, 1% (w/v) lactic acid) containing an internal standard of 1 μg/mL ergotamine-hemitartrate (1 μg/mL; Sigma) and homoperamine nitrate (2 μg/mL; custom synthesis, AgResearch). The sample was agitated in a cell disrupter (FP 120 Savant FastPrep, BIO 101 Inc.) for 20 s, speed 5, then mixed by rotation for 1 h. The extract was incubated at 4°C overnight, then centrifuged for 10 min at 3,000 g.

2.11.3 Ergovaline analysis

An 8-20 μL sample was separated by reverse phase HPLC at 28°C using a C-18 reverse phase column (5 μm, 150 x 4.6 mm, Prodigy ODS(3), Phenomenex, CA), fitted with a guard column. The mobile phase, at a flow rate of 1 mL/min, was a multilinear binary gradient separation consisting of solvent A (acetonitrile, aqueous 0.1 M ammonium acetate; 1:3 v/v) and solvent B (acetonitrile; aqueous 0.1 M ammonium acetate; 3:1 v/v) in ratios of A:B at programmed time points as follows: 0 min, 95:5; 20 min, 80:20; 35 min, 50:50; 40 min, 30:70; 45 min, 30:70; 47 min, 0:100; 55-58 min, 95:5. Ergot alkaloids were detected by fluorescence (RF551 detector, Shimadzu) with excitation at 310 nm and emission at 410 nm. The amount of ergovaline together with its natural isomer ergovalinine was estimated by comparison of peak areas (Class-LC10 software, Shimadzu) to that of the sum of internal standard ergotamine plus ergotaminine adjusted for relative fluorescence and extraction efficiency.

2.11.4 Peramine analysis

Chromatography of peramine was carried out by a column-switching procedure with an initial step to remove interfering UV absorbing compounds (Spiering, et al., 2002; Cox, Stout, 1987). Samples (50 μL) were loaded onto a RP-C8 (Alltech Associates) silica-based mixed-mode cation-exchange cartridge (7.5 x 4.6 mm, 5 μm) in a mobile phase of isopropanol:water:ammonium hydroxide (25%), 60:40:1, at a flow rate of 0.4 mL/min for 2 min. The cartridge was then flushed at 1 mL/min with 50 mM ammonium acetate, 5 mM guanidinium carbonate and 0.2 % acetic acid in water-methanol, 4:1, and analysed by normal phase HPLC using a Phenosphere (Phenomenex) silica column (250 x 4.6 mm, 5 μm) and the same solvent used for flushing. Eluted products were analysed by UV spectrophotometry at 286 nm with a UV-p70 detector (Jasco Corp.). The amount of peramine was estimated by comparison of the integrated peak area (Class-LC software, Shimadzu) of the analyte with the homoperamine internal standard.

2.11.5 Paxilline

A *P. paxilli* spore suspension was used to inoculate 25 ml of CD-YE (Section 2.2.2), trace elements (Section 2.2.7) with ~5 x 10^6 spores. The cultures were grown at 28°C for 7 days with a shaking speed of 250 rpm. The mycelia were washed three times with H_2O and freeze dried. A mycelial sample of ~0.2-0.5 g was extracted with 10 mL chloroform using a Virtis homogeniser, then 5 mL of methanol was added and the sample mixed for approximately 1 h by rotation on a spinning wheel. A sample of 1.5 mL was spun down in a micro-centrifuge for 10 min and then the supernatant dried down overnight in a fumehood. The dried extract was resuspended in 50 μL of chloroform:methanol, 2:1, and 10 μL aliquots spotted onto a TLC Plate (Polygram SIL G precoated sheets 10x20, MACHEREY-NAGEL Art.-Nr. 805013). The samples were separated in a running solvent of chloroform:acetone, 9:1. Paxilline was identified by spraying the plate with alcoholic Ehrlich's reagent (1% p-dimethylaminobenzaldehyde in 12% HCl and 50% ethanol) and heating the plate to 120°C for approximately 10 min. The migration of paxilline from *P. paxilli* strains was compared to that of authentic paxilline (AgResearch, Ruakura).

CHAPTER THREE

RESULTS

3.1 *N. lolii* and *E. festucae* contain two GGPP synthases

To isolate *ggs* sequences from the lolitrem B producing *N. lolii* Lp19 isolate, degenerate PCR was used. Lp19 was initially chosen for this study because the genome is haploid, there was a genomic library available and a supply of endophyte infected perennial ryegrass that would all be requirements for downstream applications.

The protein sequences of the available fungal GGPP synthase genes from *Neurospora crassa al-3* (Barbato, et al., 1996), *Saccharomyces cerevisiae Bts1* (Jiang, et al., 1995), *Penicillium paxilli paxG* (Young, et al., 2001), *Fusarium fujikuroi ggs-1* (Mende, et al., 1997) and *ggs-2* (Tudzynski, Hölter, 1998) were aligned to identify conserved domains, as defined by Chen, et al., (1994), that would be suitable for degenerate primer design. Primers, ggpps27, ggpps28 and ggpps29 (Table 2.3, Fig. 3.1, and Appendix 5.1), were designed to three highly conserved regions taking into consideration the placement of known introns (Fig 3.1). Degenerate PCR amplification was performed using primer pairs ggpps27 and ggpps28, and ggpps27 and ggpps29 with genomic DNA from two lolitrem-producing strains, *N. lolii* Lp19 and *E. festucae* Fl1, and a lolitrem-negative strain, *E. typhina* E8. *P. paxilli* strains PN2013 and LM662 were used as positive controls for the amplification of *ggs* genes. The *P. paxilli* wild-type strain, PN2013, contains two *ggs* genes, *paxG* and *Ppggs1*; however, these two genes can be distinguished by size, as *paxG* contains an intron within the amplified regions (Fig. 3.1). The *P. paxilli* strain LM662 is a paxilline-negative deletion mutant lacking a 22.3-kb region that contains *paxG* (Young, et al., 2001). Therefore, using DNA from this strain as a template for PCR can only result in the amplification of *ggs1*.

The degenerate PCR (Section 2.8.2) was performed with two annealing temperatures of 45°C and 47°C. Amplification of endophyte genomic DNA with primer combinations ggpps27 and ggpps28, and ggpps27 and ggpps29 resulted in 209-bp and 272-bp fragments respectively (Fig. 3.2). The PCR reactions from each primer pair were pooled and the Lp19 fragments cloned into pGEM-T (Section 2.4.4). One clone, pCY29, from primer combination ggpps27 and ggpps29, was sequenced. BLASTX analysis of the CY29 sequence showed significant similarity (E-value of 7e-41) to the *N. crassa* GGPP synthase (accession number AAC13867, locus identification number NCU01427) and other GGPP synthase sequences.

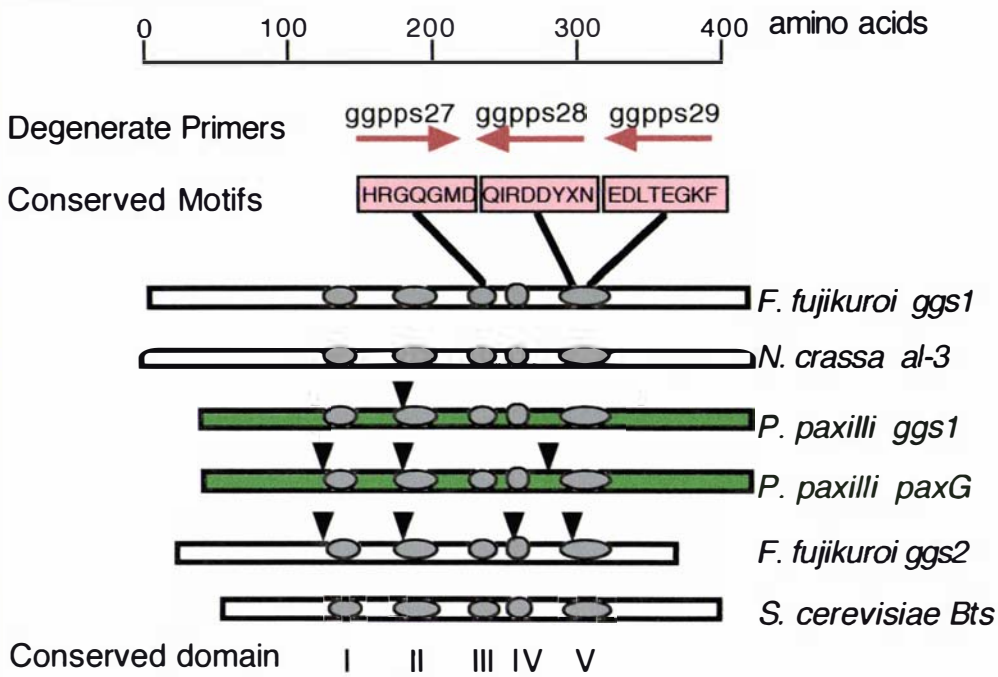


Figure 3.1 Schematic diagram of fungal geranylgeranyl diphosphate synthase sequences used for degenerate primer design

Schematic diagram of fungal geranylgeranyl diphosphate synthase sequences. The protein sequence alignment can be seen in Appendix 5.1. The rectangular boxes represent the polypeptide sequences of each gene. The grey ovals contained within each sequence represents the five conserved domains as defined by Chen et al. (1994). The primers, ggpps27, ggpps28 and ggpps29, were designed to the sequences from the highly conserved motifs, detailed in the red boxes, found in domains III and V. The primer sequences are shown in Table 2.3. The positions of the introns are represented by an arrow head. The ggs sequences from *P. paxilli* are highlighted in green.

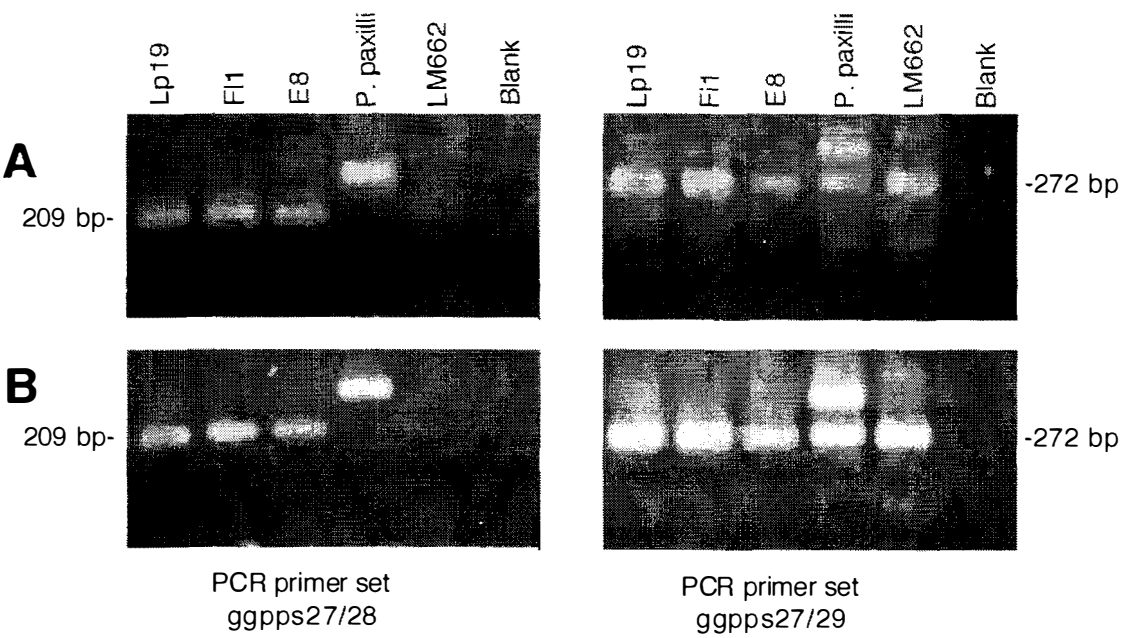


Figure 3.2 Degenerate PCR to isolate ggs fragments

PCR amplification using degenerate primers designed to *ggs* sequences with genomic DNA from *N. lolii* (Lp19), *E. festucae* (Fl1) and *E. typhina* (E8) endophyte strains, with wild-type *P. paxilli* and LM662, a *P. paxilli pax* deletion mutant used as controls. (**A**) PCR amplifications were carried out with an annealing temperature of 45^{O}C. (**B**) PCR amplifications were carried out with an annealing temperature of 47^{O}C.

The following PCR reaction conditions were used: 5 ng genomic DNA, 1x *Taq* polymerase buffer (Roche), 50 μM each dNTP, 800 nM each primer, 0.5 U *Taq* polymerase in a reaction volume of 25 μL. The PCR amplification conditions were as follows: 94^{O}C for 2 min; followed by 30 cycles of 94^{O}C for 30 s, either 45^{O}C (A) or 47^{O}C (B) for 30 s, 72^{O}C for 1 min; then one cycle of 72^{O}C for 5 min.

The PCR reactions with *P. paxilli* wild-type DNA clearly showed amplification of two bands with the PCR primer set ggpps27 and ggpps29 (Fig 3.2). However, PCR with the endophyte strains consistently produced one strong band for each primer pair. To determine whether each PCR band contained more than one product, an RFLP (restriction fragment length polymorphism) approach was used. Sequence analysis of the insert in pCY29 revealed the presence of two restriction enzyme recognition sites, *Hae*III and *Not*I (Fig 3.3) that would be unlikely to have an identical restriction endonuclease banding pattern in the sequence of a second independent *ggs* fragment. The remaining cloned fragments from both primer sets were amplified with primers ggpps27 and ggpps28 using the degenerate PCR conditions, then distinguished using RFLP analysis. The resulting fragments were digested with an appropriate enzyme (*Hae*III or *Not*I) and resolved on a 2% agarose gel. The RFLP screen on the remaining clones revealed a clone pCY28 that did not contain either the *Hae*III or *Not*I restriction-enzyme sites within the cloned fragment (Fig. 3.3). The sequence of the pCY28 cloned fragment shows similarity to *ggs* genes with the highest BLASTX match (E-value of 5e-19) to *P. paxilli Ppggs1* (accession number AAK11525; Young, et al., 2001). The pCY28 clone was identified from the PCR pool amplified with ggpps27 and ggpps28 and is therefore a shorter product than pCY29. The sequences of fragments CY28 and CY29 share 61.7% identity to each other at the nucleotide level.

To determine which clone had a partial *ggs* gene involved in lolitrem biosynthesis, each fragment was hybridised to genomic DNA from the two lolitrem-producing strains, *N. lolii* Lp19 and *E. festucae* Fl1, and the lolitrem non-producing E8 strain of *E. typhina*. The hybridising patterns (Fig. 3.4) showed that CY29 hybridised to all three strains while CY28 hybridised to just the two lolitrem producers, Lp19 and Fl1. These data support the hypothesis that CY29 is the *P. paxilli ggs1* orthologue, while CY28, which correlates with lolitrem production, is the likely *paxG* orthologue. These genes were named *ggs1* and *ltmG* (for lolitrem) respectively. Faint cross hybridisation of the CY28 fragment to the *E. typhina ggs1* is seen in the E8 lanes and corresponds to identical hybridisation patterns seen with the CY29 probe.

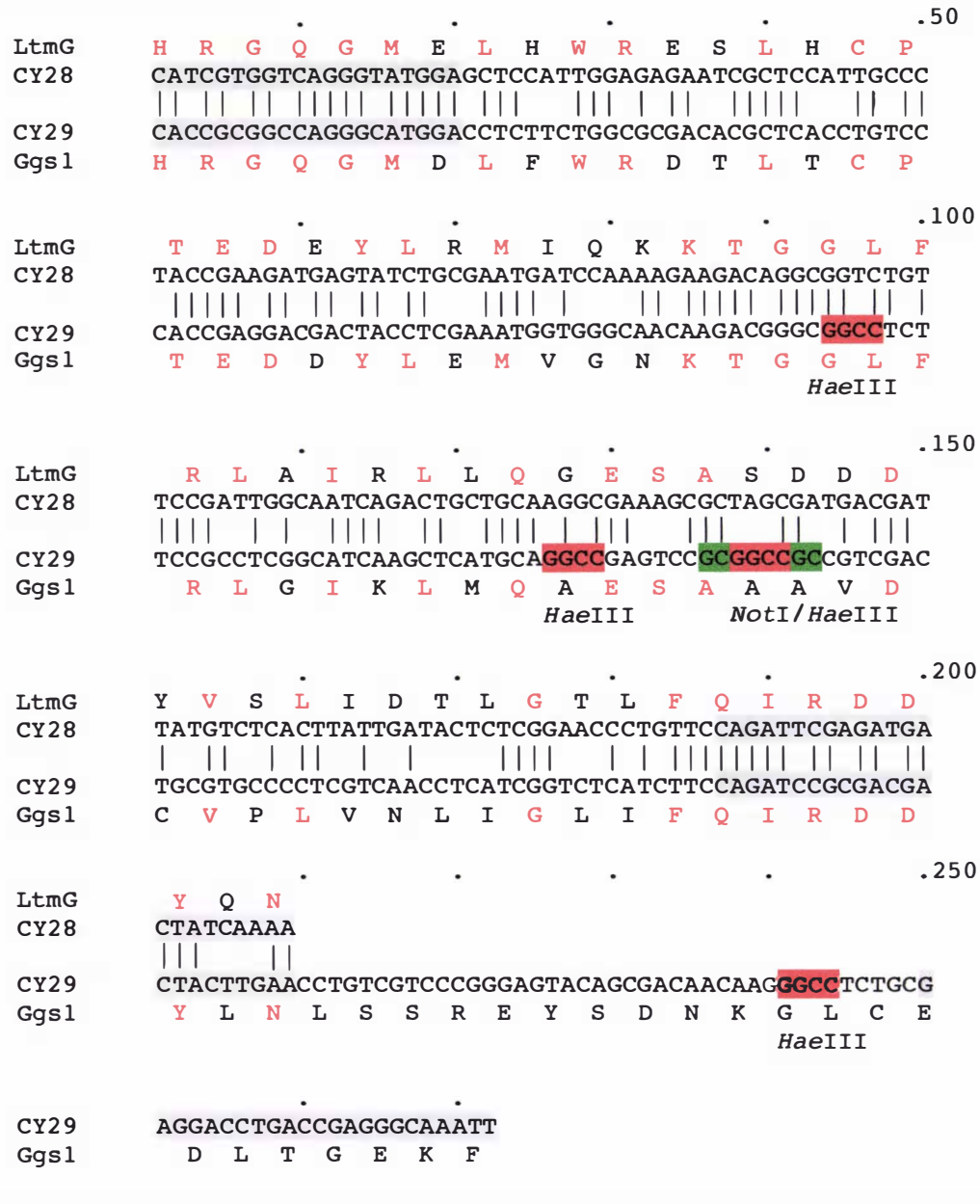


Figure 3.3 Sequence of the two isolated *ggs* fragments

The CY28 fragment was amplified with primers ggpps27 and ggpps28. The CY29 fragment was amplified with ggpps27 and ggpps29. The primer sequences are highlighted in grey. The *Hae*III and *Not*I sites are highlighted in red and green/red respectively. Sequence identity is shown by the I symbol between the two DNA sequences. The amino acid sequence identity between the two sequences is shown in red.

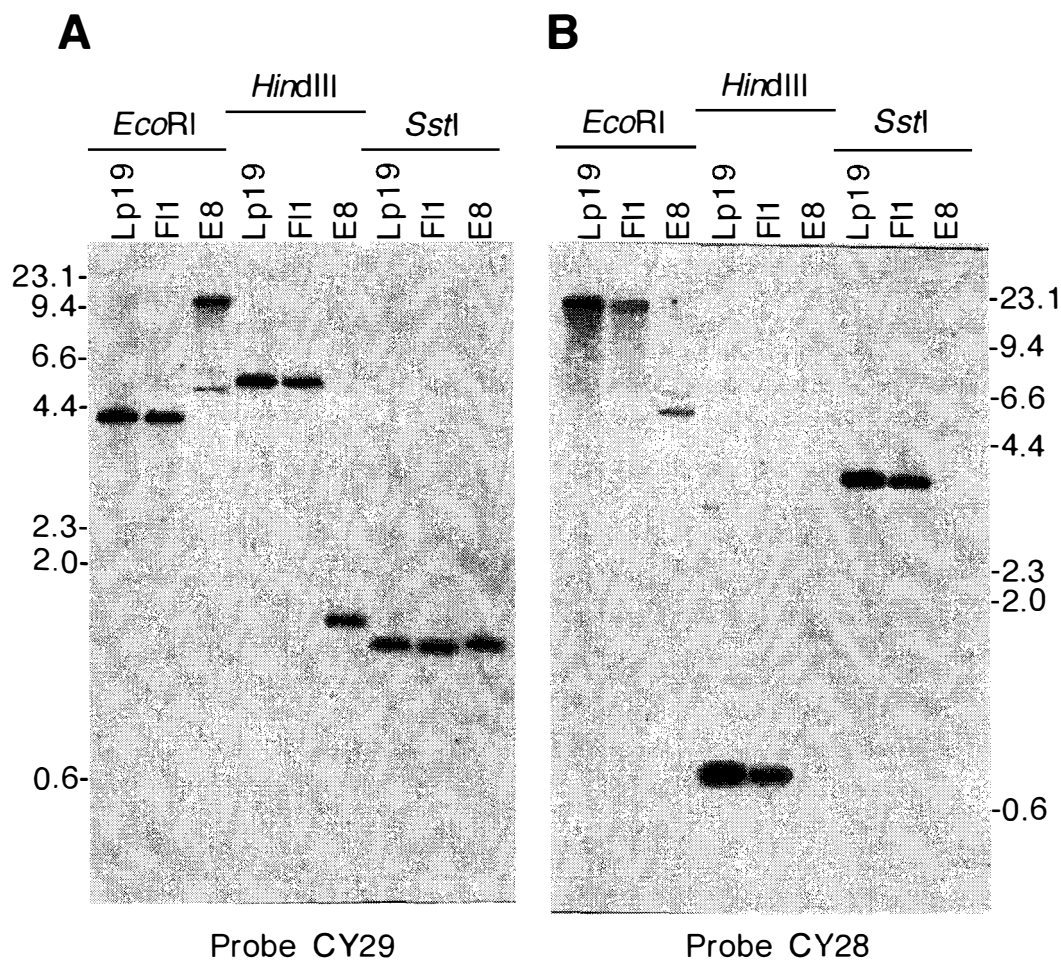


Figure 3.4 Southern analysis with the two *ggs* DNA fragments

Autoradiographs of Southern analysis of *N. lolii* Lp19, *E. festucae* Fl1 and *E. typhina* E8 hybridised with ^{32}P-labelled *ggs* fragments (**A**) CY29 (*ggs1*) and (**B**) CY28 (*ltmG*). The DNA was digested with the restriction enzymes *Eco*RI, *Hin*dIII and *Sst*I. Faint cross hybridisation of CY28 to the E8 *ggs1* gene is seen in B. The size standards are in kb.

3.2 The *ggs1* gene

3.2.1 The *N. lolii ggs1* gene

The complete *N. lolii ggs1* gene was isolated from an Lp19 λGEM12 genomic library (Section 2.5) using the fragment CY29 as a probe. The CY29 fragment was well represented within the library with ~100 hybridising plaques from ~100,000 screened. Of the 13 clones initially selected, 11 remained positive through three rounds of library screening used to purify the clones. A 4.2-kb region was sequenced from λCY100 and λCY102 using a combination of primer walking and sequencing of fragments cloned from *Sst*I digests of λCY100 and λCY102.

The *ggs1* gene was contained within a single ORF as shown in Figure 3.5. The 5' region and putative start site of this gene were difficult to characterise as this is the least conserved region of GGPP synthases and *ggs1* contains several repetitive elements within this region. There are five possible in-frame start codons. However, the second methionine has the best Kozak sequence, which is 100% identical to the *N. crassa* consensus of $CN_3CAMVATGGC$ (Bruchez, et al., 1993). Sequence analysis of *ggs1* predicts the gene lacks introns, in particular, the single intron found in *Ppggs1* and the two conserved introns found in *P. paxilli paxG*, *A. flavus atmG* and *ggs-2* from *F. fujikuroi*. A polypeptide motif (IPPRXSS) common to the N-terminal region of *N. crassa*, *P. paxilli*, *A. nidulans*, *F. fujikuroi* and *M. grisea ggs1* genes, is also present in the longer proposed transcript of *N. lolii ggs1*, starting at the second methionine (Fig. 3.5, Appendix 5.1). Three repetitive DNA sequences, $CAR_{(10)}$, $GCYATCAAR_{(6)}$ and $CAG_{(12)}$ do not appear to be contained in introns (Fig 3.5). The deduced Lp19 Ggs1 encodes a polypeptide of 466 residues with a predicted unmodified mass of 51.9 kDa. BLASTP analysis of the *N. lolii* Ggs1 showed it was most similar to that of *Gibberella zeae* (accession number AEE68323, *F. graminearum* FG10097 locus) with an E-value of e-136. The BLASTP analysis of *N. lolii* Ggs1 to the *P. paxilli* genes, PpGgs1 and PaxG, yields E-values of 1e-106 and 2e-79 respectively.

The chromosomal location of the *ggs1* gene was determined by Southern analysis of chromosomal DNA separated by pulse field gel electrophoresis and then hybridised with a *ggs1* fragment (Fig. 3.6). The Lp19 *ggs1* gene is contained on a ~6.2 Mb

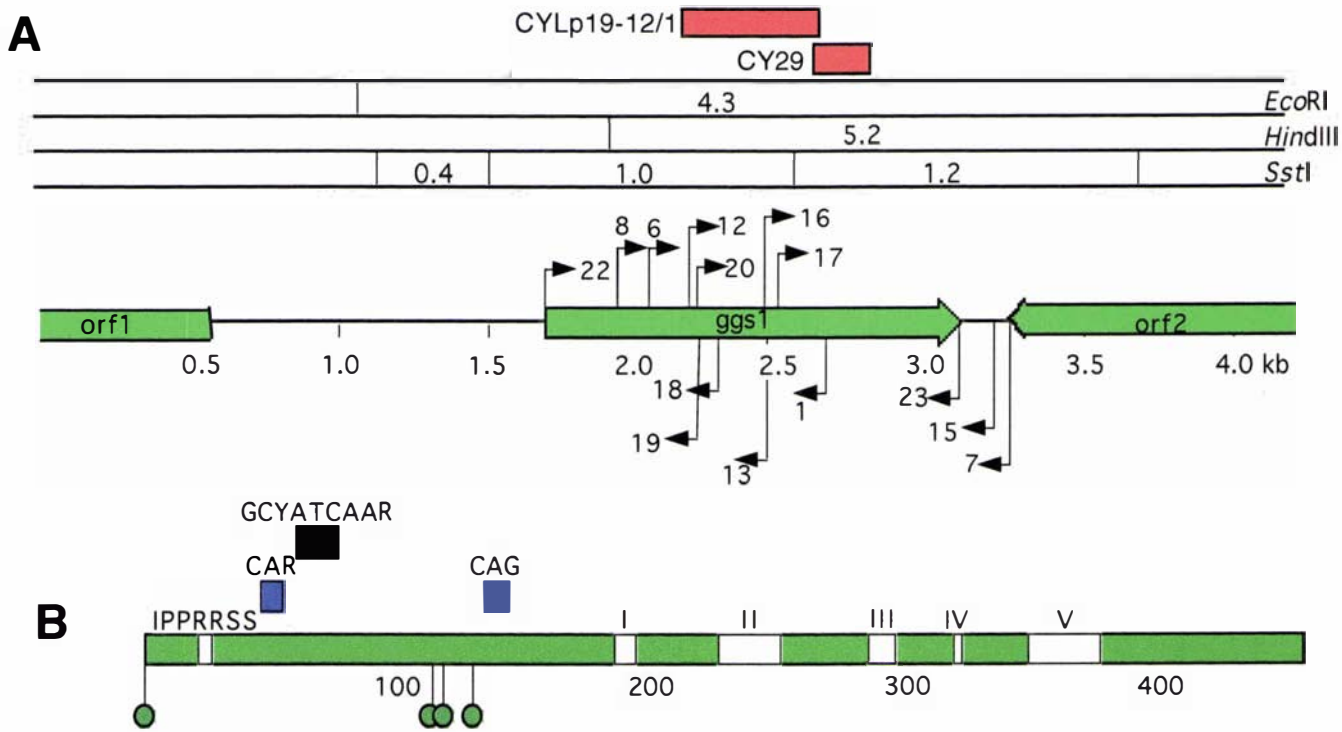


Figure 3.5 Restriction enzyme map of the *ggs1* gene and flanking region

(A) Restriction enzyme map of the *ggs1* region. The red boxes show the region of the hybridisation probes, the genes are shown as green arrows, the position and orientation of the primers are shown as black arrows. The primer names have been abbreviated from CYLp19-# to just the number.

(B) A schematic diagram of the *ggs1* polypeptide. The five conserved domains and the IPPRRSS domain are shown as white boxes on the green polypeptide. The position of four of the five possible start codons are green lollypops and the location of the repetitive DNA motifs are blue boxes.

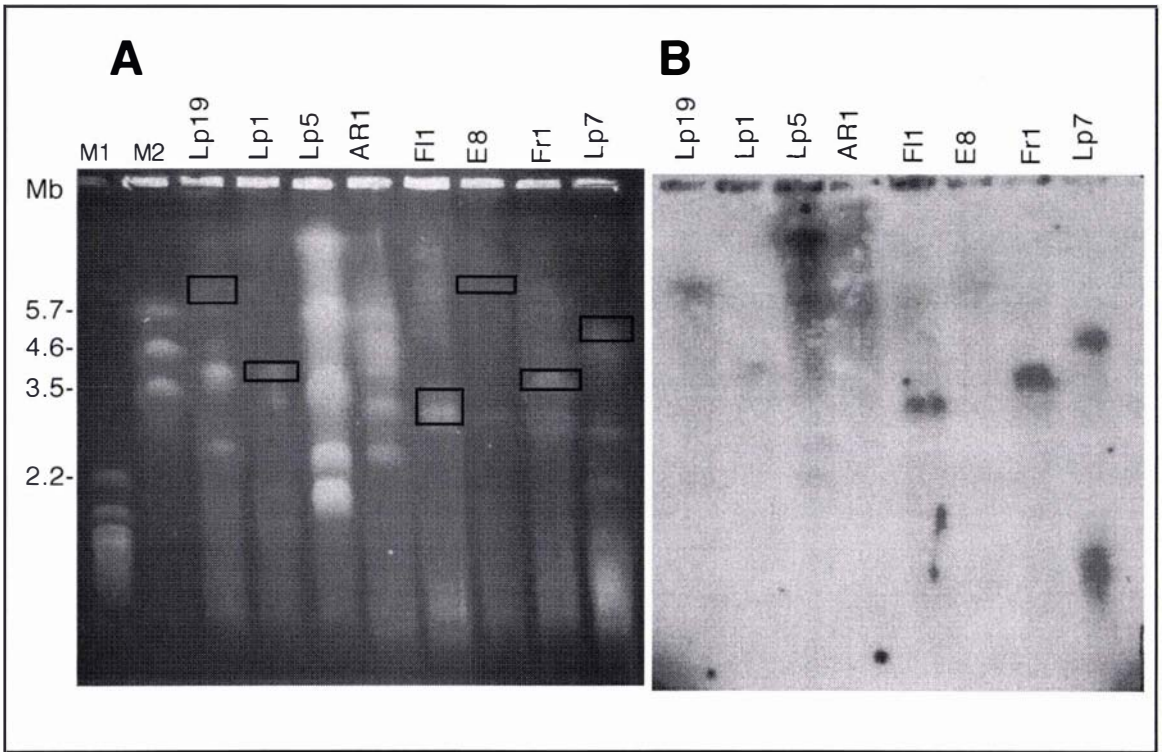


Figure 3.6 The chromosomal location of the *ggs1* gene

The chromosomal location of the *ggs1* gene. (**A**) Chromosomal DNA separation of *N. lolii* Lp19, Lp5, AR1 and Lp7, *E. festucae* Fl1, Fr1, *E. typhina* E8 and *Neotyphodium* sp. LpTG-2 Lp1. The size standards are indicated in Mb based on the mobility of the markers, M1, *S. cerevisiae* (BioRad) and M2, *S. pombe* (BioRad). The electrophoresis conditions were (Kuldau et al, 1999) 0.6% chromosomal grade agarose (BioRad) in 0.5x TBE at 40 volts. The gel was run at 10^{0}C with switch times ramped from 1200-6000 s for 240 h. (**B**) Autoradiograph of the gel from **A** hybridised with a ^{32}P-labelled *ggs1* fragment amplifed with primers CYLp19-12 and CYLp19-1 (the location of the probe is shown in Fig. 3.5). The hybridising chromosomal DNA bands are surrounded by a black box in panel A. The Lp5 and AR1 hybridising chromosomal DNA bands have not been marked due to the quality of the hybridisation. Two hybridising bands are present in the Lp7 chromosomal DNA separation but the low molecular weight fragment is likely to be due to degradation.

chromosome, which is a similar size to that of E8. Of the remaining strains that were screened, none appear to have the *ggs1* probe hybridising to the same size chromosome (Fig. 3.6).

Adjacent to the *ggs1* gene are two sequences of partial genes, *orf1* and *orf2* (Fig. 3.5). BLASTP analysis of Orf1 and Orf2 showed that these sequences were most similar to *F. graminearum* FG10100 and FG10096, respectively (Table 3.1). However, these gene products have no known function. The three Lp19 genes (*orf1*, *orf2* and *ggs1*) were used to search the *M. grisea*, *N. crassa*, *A. nidulans* and *F. graminearum* genome sequences (http://www.broad.mit.edu/annotation/fungi/fgi) for homologues, and then to determine if the genomic arrangement of these genes was conserved across species. The BLASTP analysis of the three *N. lolii* genes to those of the four species mentioned above is shown in Table 3.1. The genetic map of the *N. lolii ggs1* locus was compared to those of *M. grisea*, *N. crassa*, *A. nidulans* and *F. graminearum* looking for possible microsynteny (Fig. 3.7, Table 3.1). Of these four fungal species, *M. grisea* has a genetic arrangement that is most similar to *N. lolii* with respect to both the gene order and orientation as these three genes are found in an *M. grisea* contig (Fig. 3.7, Table 3.1). The genes in *F. graminearum* are in the same orientation as in Lp19 (Fig. 3.7, Table 3.1). However, there are two additional genes, FG10098 and FG10099, inserted between the *F. graminearum* equivalent to *N. lolii orf1* and *ggs1*. *N. crassa* has the equivalent of *orf1* immediately adjacent to *orf2* but *orf1* is in the opposite orientation from that of the Lp19 gene and there is no intervening *ggs1* equivalent at this locus (Fig. 3.7). The *N. crassa orf1* and *orf2* genes are contained in contig 3.384 in supercontig 30, while *al-3*, the *ggs* gene, is contained in a separate contig of 3.58 in supercontig 3 (Table 3.1). AN7677, an *A. nidulans* sequence most similar to *orf2*, is contained in contig 1.130 in supercontig 10, whereas *orf1* and *ggs1* equivalents are contained in contig 1.10 in supercontig 1 (Table 3.1). The distances between the genes are variable when comparing the *N. lolii* sequence and each of the four fungal species mentioned above. The three genes AN0653 from *A. nidulans* and the *F. graminearum* genes FG10099 and FG10098 (Fig. 3.7), associated with *orf1* and *ggs1* from *A. nidulans* and *F. graminearum* respectively, can be found within the other completed genomes when using the BLASTP and/or tBLASTN algorithms, with only *N. crassa* lacking a FG10099 homologue. However, none of these homologues is located in the same contig, with only MG00364, the homologue of AN0653, located in the same super-contig as those clustered with *ggs1*.

Table 3.1 Identification of orthologues to *N. lolii ggs1*, *orf1* and *orf2*

Gene	Putative function[1]	ID	Contig	Super-contig	E-Value[2]	species
orf1[3]	Unknown	MG00757	2.134	1	6e-33	*M. grisea*
		NCU06636	3.384	30	5e-43	*N. crassa*
		AN0655	1.10	1	7e-34	*A. nidulans*
		FG10100	1.418	7	5e-44	*F. graminearum*
ggs1	geranylgeranyl diphosphate synthase	MG00758	2.134	1	e-127	*M. grisea*
		NCU01427	3.57	3	e-121	*N. crassa*
		AN0654	1.10	1	e-104	*A. nidulans*
		FG10097	1.418	7	e-136	*F. graminearum*
orf2[3]	Unknown	MG00759	2.134	1	4e-74	*M. grisea*
		NCU06637	3.384	30	3e-74	*N. crassa*
		AN7677	1.130	10	9e-49	*A. nidulans*
		FG10096	1.418	7	1e-98	*F. graminearum*

[1]Putative function is predicted, based on the best BLASTP match to a protein with a predicted function or homology.

[2]Homology between the three *N. lolii* genes and four fungal species stated is based on BLASTP.
The searches were completed through http://www.broad.mit.edu/annotation/fungi/fgi

[3]The Lp19 Orf1 and Orf2 polypeptide sequences used in the BLASTP search are partial sequences and would potentially have a higher E-value if complete.

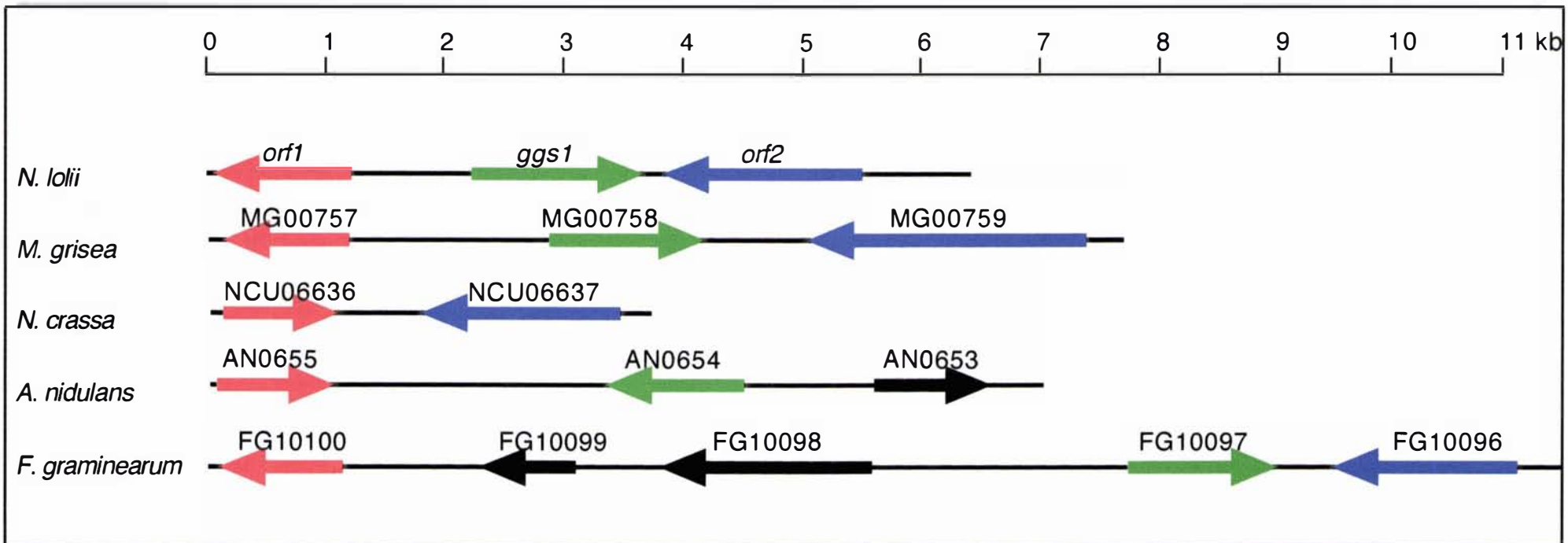


Figure 3.7 Schematic diagram showing microsynteny of the *ggs1* locus between the fungal species, *N. lolii*, *M. grisea*, *N. crassa*, *A. nidulans* and *F. graminearum*

The order and orientation of *N. lolii* genes clustered with the *ggs1* gene as seen in the fungal strains *M. grisea*, *N. crassa*, *A. nidulans* and *F. graminearum*. The red arrow is *orf1*, or genes that are homologous to *orf1*, the green arrow is *ggs1* or genes homologous to *ggs1*, the blue arrow is *orf2* or genes homologous to *orf2* and the black arrows are genes that are unrelated to the current *N. lolii* organisation. Genes from *M. grisea*, *N. crassa*, *A. nidulans* and *F. graminearum* are identifiable by their locus number as per http://www.broad.mit.edu/annotation/fungi/fgi. The *ggs1* gene (*al-3*) from *N. crassa* is located on a separate contig from *orf1* and *orf2*.

3.2.2 The *E. festucae* and *E. typhina ggs1* gene

To confirm the 5' region and the putative translational start site, the *ggs1* orthologues were isolated from *E. festucae* strain Fl1 and *E. typhina* strain E8 using PCR and IPCR with primers designed to the *N. lolii*, Lp19 sequence (Fig. 3.8). Sequence analysis of the *thi* and *prt1* genes between Lp19, Fl1, and E8 has shown a greater nucleotide sequence variation within the promoter regions than within the coding regions (Zhang, Bryant and Scott unpublished). To isolate the 5' region of the *ggs1* gene IPCR was used as Southern analysis revealed that the *ggs1* fragment hybridised to a 1 kb *Sst*I band in all three strains (Fig. 3.8). The primers, CYLp19-18 and CYLp19-16, were used to amplify a ~860 bp fragment with *Sst*I-digested then self-ligated genomic DNA (Fig. 3.8C) and this PCR product was purified then sequenced with CYLp19-19 and CYLp19-17. The 3' region of the *ggs1* gene including the 3' untranslated region was amplified and sequenced with primers CYLp19-20 and either CYLp19-15 or CYLp19-7 (Fig 3.8D). Based on these data, complete sequences of the *ggs1* genes from Fl1 and E8 were assembled (Appendix 5.3).

The DNA sequence of all three strains contained the five putative ATG start sites and have the repetitive elements CAR, GCYATCAAR and CAG (Fig. 3.9). However, there is no evidence of consistent intron splice sites among the three endophyte strains that would result in excision of the repeated sequences mentioned above. The proposed transcripts of *Nlggs1* and *Efggs1* share 99.6% identity at the nucleotide level, whereas the *Etggs1* has diverged sharing 92% identity to *Nlggs1* and *Efggs1*. The E8 polypeptide sequence, when compared to Lp19 and F1, contains just six amino acid replacements within the most highly conserved region of the gene starting beyond the CAG repeat (Appendix 5.1). All six changes are conservative.

The repetitive DNA sequences are polymorphic between the three strains with the CAG repeat sequence unique to each strain. To determine whether the polymorphic CAG repeat would be a useful marker sequence for *Neotyphodium* and *Epichloe* strain identification, this region was amplified from a number of endophyte isolates using the primers CYLp19-6 and CYLp19-1 (Fig. 3.5). The resulting PCR products were purified (Section 2.4.3) and sequenced with the primer CYLp19-13. Sequence was obtained from the majority of the endophyte isolates screened; however, those known to be hybrids contained more than one sequence and were not analysed further. The Lp1

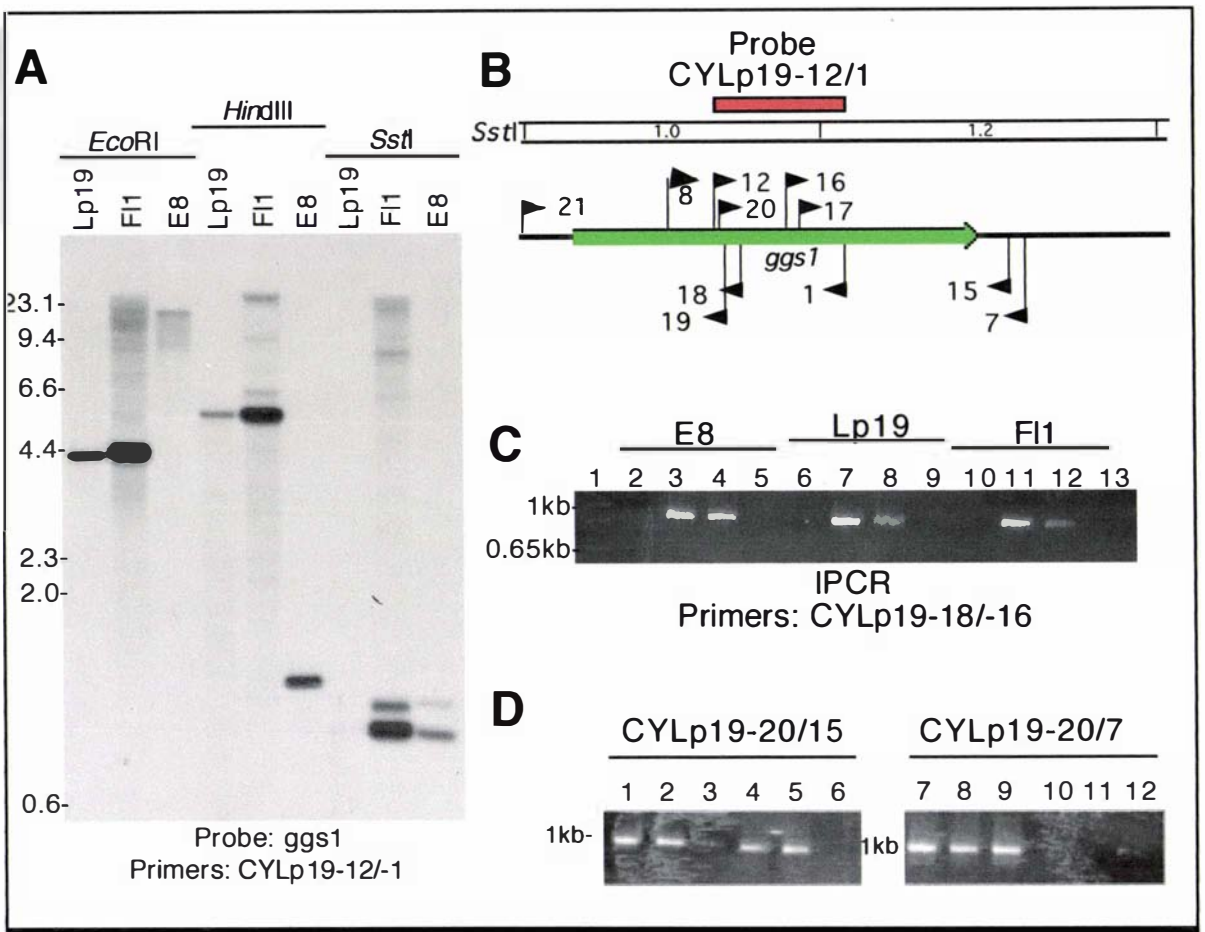


Figure 3.8 Isolation of *ggs1* from *E. festucae* and *E typhina*

(**A**) Autoradiograph of Southern analysis of *N. lolii* Lp19, *E. festucae* Fl1, and *E. typhina* E8 genomic DNA digested with *Eco*RI, *Hin*dIII and *Sst*I and hybridised with a ^{32}P-labelled *ggs1* probe as shown in B. Note: The DNA in the Lp19-*Sst*I lane was degraded, but the nucleotide sequence data confirmed the expected fragment size.

(**B**) A genomic map of the Lp19 *Sst*I fragments, based on nucleotide sequence, that hybridised to the *ggs1* probe, amplified with primers CYLp19-12 and CYLp19-1 and the primer locations. The primer names have been shortened from CYLp19-# to just the #.

(**C**) Isolation of the 5' region of *ggs1* from Lp19, Fl1 and E8. IPCR with *Sst*I-digested then self-ligated genomic DNA amplified with primers CYLp19-18 and CYLp19-16. Lane (1) 1kb+ ladder (Invitrogen); (2) E8 genomic DNA uncut; (3 and 4) E8 genomic DNA, *Sst*I-digested then self-ligated; (5, 9 and 13) water control; (6) Lp19 genomic DNA uncut; (7 and 8) Lp19 genomic DNA, *Sst*I-digested then self-ligated; (10) Fl1 genomic DNA uncut; (11 and 12) Fl1 genomic DNA, *Sst*I-digested then self-ligated.

(**D**) Isolation of the 3' region of *ggs1* from Lp19, Fl1 and E8. Lane (1 and 7) Lp19; (2, 3, 8 and 9) Fl1; (4, 5, 10 and 11) E8; (6 and 12) water control.

The following PCR reaction conditions were used: 50 ng SstI-digested/self-ligated genomic DNA (C) or 5 ng genomic DNA (D), 1x *Taq* polymerase buffer (Roche), 50 μM each dNTP, 200 nM each primer, 0.5 U *Taq* polymerase in a reaction volume of 25 μL. (C) The PCR amplification conditions were as follows: 94^{o}C for 2 min; followed by 30 cycles of 94^{o}C for 15 s, 60^{o}C for 30 s, 72^{o}C for 1 min; then one cycle of 72^{o}C for 10 min. (D) The PCR amplification conditions were as follows: 94^{o}C for 2 min; followed by 30 cycles of 94^{o}C for 15 s, 55^{o}C for 30 s, 72^{o}C for 1 min 20 s; then one cycle of 72^{o}C for 10 min.

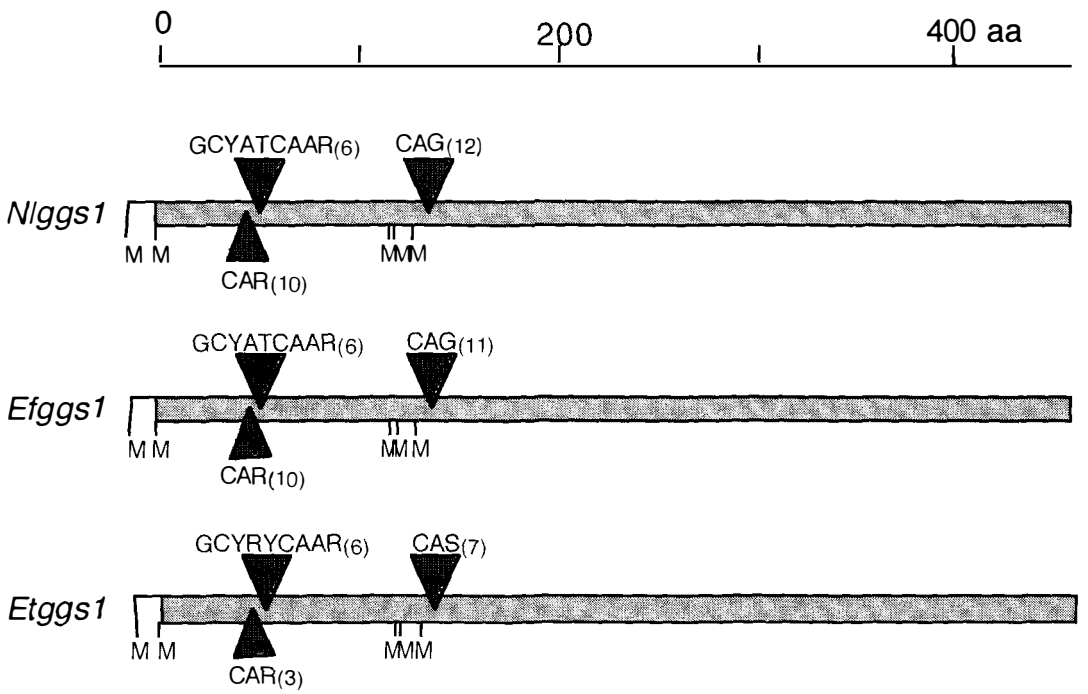


Figure 3.9 Schematic diagram of the endophyte *ggs1* genes

A schematic diagram of the *ggs1* genes from *N. lolii* Lp19, *Nlggs1*; *E. festucae* Fl1, *Efggs1;* and *E. typhina* E8, *Etggs1*. The grey rectangle indicates the coding region of the gene starting at the second proposed methionine. The position of the repetitive elements are shown as triangles and are refered to as their DNA sequence. The positions of the five possible in-frame start methionines are shown as an M. The approximate size of the polypeptides is represented as amino acid residues. The following ambiguity code has been used in the repetitive elements; Y = C or T; R = A or G; S = G or C.

strain is the only example of a hybrid where at least one of the ancestral sequences was isolated. This was achieved using the primer CYLp19-8 (in place of CYLp19-6), which is mismatched with the *E. typhina* sequence and therefore amplifies only the *N. lolii* sequence. The generated sequences are aligned in Table 3.2. The most obvious pattern identified was that for the clade containing the *E. typhina* complex, where CAG is a CAS sequence, where S could be either a G or C nucleotide. The CAG repeat length does not correlate with an indole-diterpenoid chemotype.

3.2.3 Expression of *ggs1*

Expression analysis of the *ggs1* gene by RT-PCR was determined for *N. lolii* Lp19 and *E. festucae* Fl1 and used to predict the proposed translational start site of *ggs1*. RNA was isolated from Lp19 and Fl1 grown in liquid PD media (Section 2.2.5) and total RNA was used in a one step RT-PCR reaction (Invitrogen). Primer CYLp19-22, which anneals to the putative ATG start site, and primer CYLp19-18 were used to amplify the 5' region. A band of the same size as the PCR product from genomic DNA confirmed the absence of introns within *ggs1* (Fig. 3.10). As a control PCR reaction for the *ggs1* gene, primers CYLp19-16 and CYLp19-1, which anneal to more conserved regions of the gene, were also used. Both *ggs1* primer pairs show that the *ggs1* gene is expressed in culture and lacks introns. The *tub2* gene, amplified with primer pair T1.1 and T1.2 (Fig. 3.10), was used as a control to show there was no contaminating DNA in the RNA samples as the corresponding gene contains introns that require splicing. No products were amplified from the RNA samples subjected to a PCR only reaction, confirming that there was no DNA contamination in the RNA (Fig. 3.10).

3.3 The *ltmG* gene

The *ltmG* fragment, CY28 (Section 3.1), was used as a probe to isolate sequences from an Lp19 λGEM12 genomic library (Section 2.5). This region was under represented in the library with only five clones isolated from ~80, 000 plaques plated. The largest lambda clone, λCY218, ~15.6 kb (Fig. 3.11), was completely sequenced and shown to contain a complete copy of the *ltmG* gene.

Table 3.2 Sequence analysis across the polymorphic CAG repeat

Strain	Species	Ancestral grouping	1	2	3	4	5	14	13	12	11	10	9	8	7	6	5	4	3	2	1
E52	*E. amarillans*		CAT	CAT	CAT	CAT	...	...	...	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
E57	*E. amarillans*		CAT	CAT	CAT	...	...	...	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
E1031	*E. baconii*		CAT	CAT	CAT	...	...	...	...	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
E248	*E. baconii*		CAT	CAT	CAT	...	...	...	CAG	CAG	CAG	CAT	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
E1040	*E. brachyelytri*		CAT	CAT	CAT	...	...	...	...	...	...	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG
E799	*E. bromicola*		CAT	CAT	...	...	...	...	...	...	...	...	...	...	...	...	...	CAG	CAG	CAG	CAG
E501	*E. bromicola*		CAT	CAT	...	...	...	...	...	...	...	...	...	...	...	...	...	CAG	CAG	CAG	CAG
E56	*E. elymi*		CAT	CAT	CAT	CAT	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
E184	*E. elymi*		CAT	CAT	CAT	...	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
Frc7	*E. festucae*		CAT	CAT	CAT	CAT	CAT	...	...	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
Fg1	*E. festucae*		CAT	CAT	CAT	CAT	...	...	...	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
Frr1	*E. festucae*		CAT	CAT	CAT	CAT	...	...	...	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
Frc5	*E. festucae*		CAT	CAT	CAT	CAT	...	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
Fl1	*E. festucae*		CAT	CAT	CAT	...	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
Fr1	*E. festucae*		CAT	CAT	CAT	...	...	...	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
E189	*E. festucae*		CAT	CAT	CAT	...	...	...	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
E2772	*E. glyceriae*		CAT	CAT	CAT	...	...	...	...	...	...	...	...	...	...	CCT	CAT	CAG	CAG	CAG	CAG
E503	*E. sylvatica*	ETC	CAT	CAT	CAT	...	...	...	...	...	...	CAG	CAC	CAG	CAC	CAG	CAG	CAC	CAG	CAC	CAG
E354	*E. sylvatica*	ETC	CAT	CAT	CAT	...	...	...	...	...	...	...	...	CAG	CAC	CAG	CAG	CAC	CAG	CAC	CAG
E348	*E. typhina*	ETC	CAT	CAT	...	...	...	...	...	...	...	...	...	...	...	...	...	...	CAG	CAC	CAG
E425	*E. typhina*	ETC	CAT	CAT	CAT	..	...	...	...	...	...	CAG	CAC	CAG	CAG	CAC	CAG	CAG	CAC	CAG	CAG
E2463	*E. typhina*	ETC	CAT	CAT	CAT	...	...	...	...	CAG	CAC	CAG	CAG	CAC	CAG	CAG	CAC	CAG	CAG	CAC	CAG
E505	*E. typhina*	ETC	CAT	CAT	CAT	...	...	...	...	...	...	...	...	...	...	...	CAG	CAC	CAG	CAC	CAG
E8	*E. typhina*	ETC	CAT	CAT	CAT	...	...	...	...	...	...	...	...	...	CAG	CAC	CAG	CAC	CAG	CAC	CAG
Poa	*E. typhina*	ETC	CAT	CAT	...	...	...	CAG	CAC	CAG	CAC	CAG	CAC	CAG	CAC	CAC	CAG	CAC	CAG	CAC	CAG
E1022	*E. typhina*	ETC	CAT	CAT	...	...	...	...	...	...	...	CAG	CAC	CAC	CAC	CAC	CAC	CAC	CAG	CAG	CAG
E899	*N. aotearoae*	Nao	CAT	CAT	CAT	...	...	CAG	CAC	CAG	CAG	CAC	CAG	CAG	CAC	CAG	CAG	CAC	CAG	CAG	CAG
E818	*N. inebrians*	Nin	CAT	CAT	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	CAG	CAG
Lp9	*N. lolii*	Efe	CAT	CAT	CAT	CAT	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
Lp10	*N. lolii*	Efe	CAT	CAT	CAT	CAT	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
Lp3	*N. lolii*	Efe	CAT	CAT	CAT	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
Lp4	*N. lolii*	Efe	CAT	CAT	CAT	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
Lp5	*N. lolii*	Efe	CAT	CAT	CAT	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
Lp7	*N. lolii*	Efe	CAT	CAT	CAT	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
Lp13	*N. lolii*	Efe	CAT	CAT	CAT	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
Lp19	*N. lolii*	Efe	CAT	CAT	CAT	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
Lp14	*N. lolii*	Efe	CAT	CAT	CAT	...	...	...	...	...	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG
AR1	*N. lolii*	Efe	CAT	CAT	CAT	...	...	...	...	...	...	...	...	...	...	...	CAG	CAG	CAG	CAG	CAG
E822	*N. melicicola*	Efe/Nao	CAT	CAT	CAT	CAT	...	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG	CAG
Lp1	*LpTG-2*	Efe/ETC	CAT	...	...	...	...	...	...	...	...	...	...	...	...	CAG	CAG	CAG	CAG	CAG	CAG

The red text indicates nucleotide bases that are different from the CAG consensus
The abbreviations for the ancestral groupings are ETC = *E. typhina* complex, Nao = *Neotyphodium aotearoae*, Nin = *N. inebrians*, Efe = *E. festucae*.

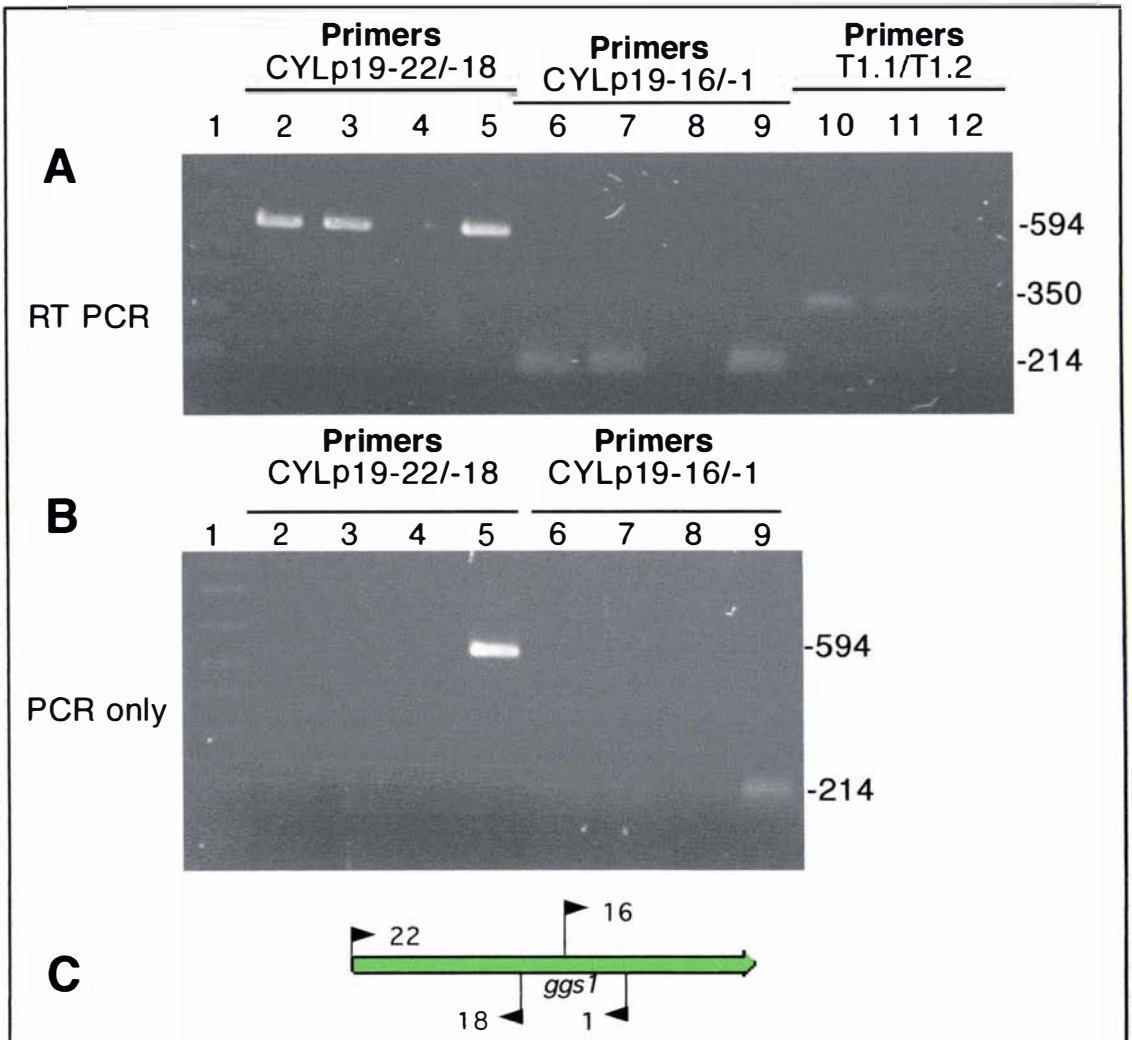


Figure 3.10 Expression analysis of the *ggs1* gene

RT-PCR analysis on the *ggs1* gene. Lanes (1) 1kb+ ladder (Invitrogen); (2, 6 and 10) Lp19 RNA; (3, 7 and 11) Fl1 RNA; (4, 8 and 12) water; (5 and 9) Lp19 genomic DNA. The fragment sizes are indicated as bp. The expected size of a tub2 genomic fragment would have been 727 bp. (**A**) RT-PCR reaction. (**B**) PCR only. All fragment sizes are indicated in bp. (**C**) schematic diagram of *ggs1* showing the location of the primers. The primer names have been shortened from CYLp19-# to just the number.

The following PCR reaction conditions were used: (A and B) 1 μg of total RNA or 5 ng of genomic DNA, 1x Superscript RT-one step (Invitrogen) reaction mix, 200 nM each primer with 0.5 μL of Superscript RT-enzyme mix in a reaction volume of 25 μL. The PCR amplification conditions were as follows: **(A)** 50^{o}C for 20 min then 94^{o}C for 2 min followed by 35 cycles of 94^{o}C for 15 s, 60^{o}C for 30 s and 72^{o}C for 1 min, then 1 cycle of 72^{o}C for 10 min. (**B**) the initial 50^{o}C step was excluded.

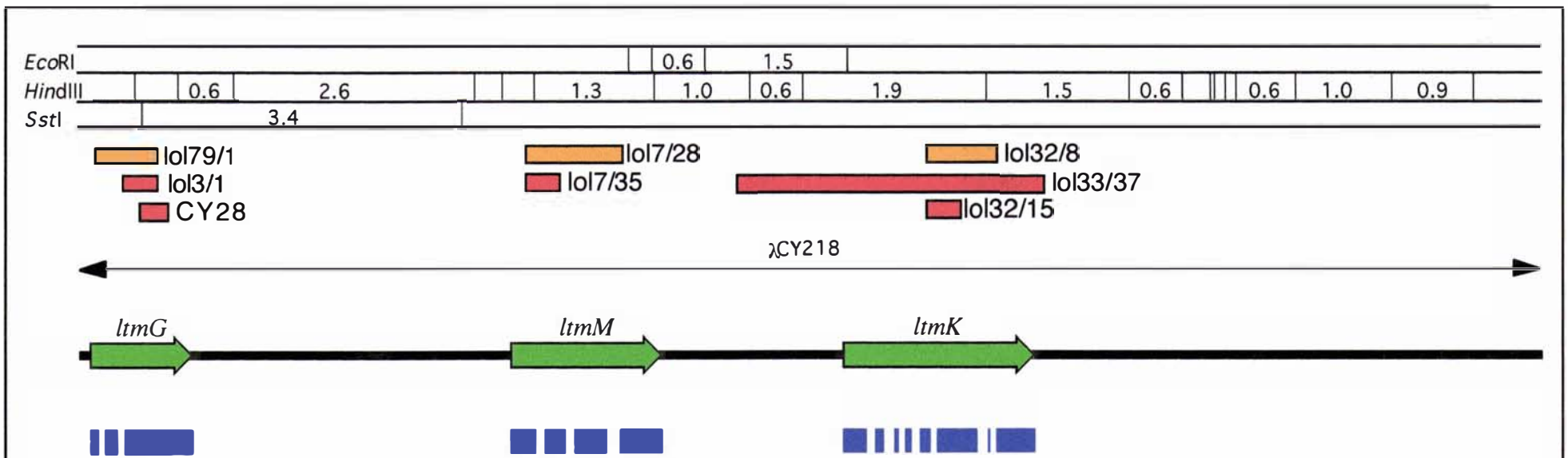


Figure 3.11 Genomic map of the λCY218 region

A restriction enzyme and gene map of the Lp19 λCY218 region. The genes are shown as green arrows while the exons are represented as blue boxes. The position of probes amplified from genomic DNA are shown as red boxes and those amplified from cDNA are shown as orange boxes, with the primer combinations used for each probe amplification are adjacent to each box.

Sequence analysis of *ltmG* predicts the presence of two introns (Fig. 3.11 and 3.12). These two introns were confirmed by RT-PCR analysis using RNA isolated from endophyte-infected ryegrass. The *ltmG*-coding region was amplified from cDNA, generated from random priming of mRNA, using the primer combination of lol79 and lol1 (Fig. 3.12A). The resulting RT-PCR product was purified and directly sequenced with gene-specific primers then compared to the genomic sequence for confirmation of the intron splice sites (Table 3.3).

The FastA analysis shows that LtmG is 54.1% identical to PpGgs1, 52.6% identical (across 302 amino acids) to PaxG and 59.4% identical (across 293 amino acids) to AtmG. The top BLASTP database match is to that of *F. graminearum* (accession number EAA72205) with an E-value of 6e-96. The Ggs1 and PaxG polypeptide sequences from *P. paxilli* have BLASTP E-values of 1e-90 and 3e-86, respectively. The *ltmG* gene contains two of the three introns found in *P. paxilli paxG* (Young, et al., 2001) and *A. flavus atmG* (Zhang, et al., 2004) and two of the four introns found in *F. fujikuroi ggs2* (Tudzynski, Hölter, 1998) (Appendix 5.1). The placement and phase of the introns is conserved.

3.4 The *ltmG* gene is contained within a gene cluster

Adjacent to *ltmG* are two genes, *ltmM* and *ltmK* (Fig. 3.11) encoding, an FAD dependent monooxygenase and P450 monooxygenase, respectively. LtmM and LtmK have significant similarity to *P. paxilli* PaxM and PaxP, respectively. Therefore, *ltmG* forms a gene cluster with homologues of *paxM* (*ltmM*) and *paxP* (*ltmK*). The AT content flanking *ltmK* is extremely high at ~70% AT and devoid of open reading frames. No other genes were evident from sequence analysis using BLAST searches.

3.4.1 *ltmM*, an FAD dependent monooxygenase

Sequence analysis of *ltmM* predicts the presence of three introns. These three introns were confirmed by RT-PCR analysis using RNA isolated from endophyte-infected perennial ryegrass (Fig. 3.12B). The *ltmM* coding region was amplified from cDNA, synthesised by random priming of mRNA, using the primer combinations of lol7 and lol35 to distinguish intron one, and primers lol14 and lol28 for introns two and three

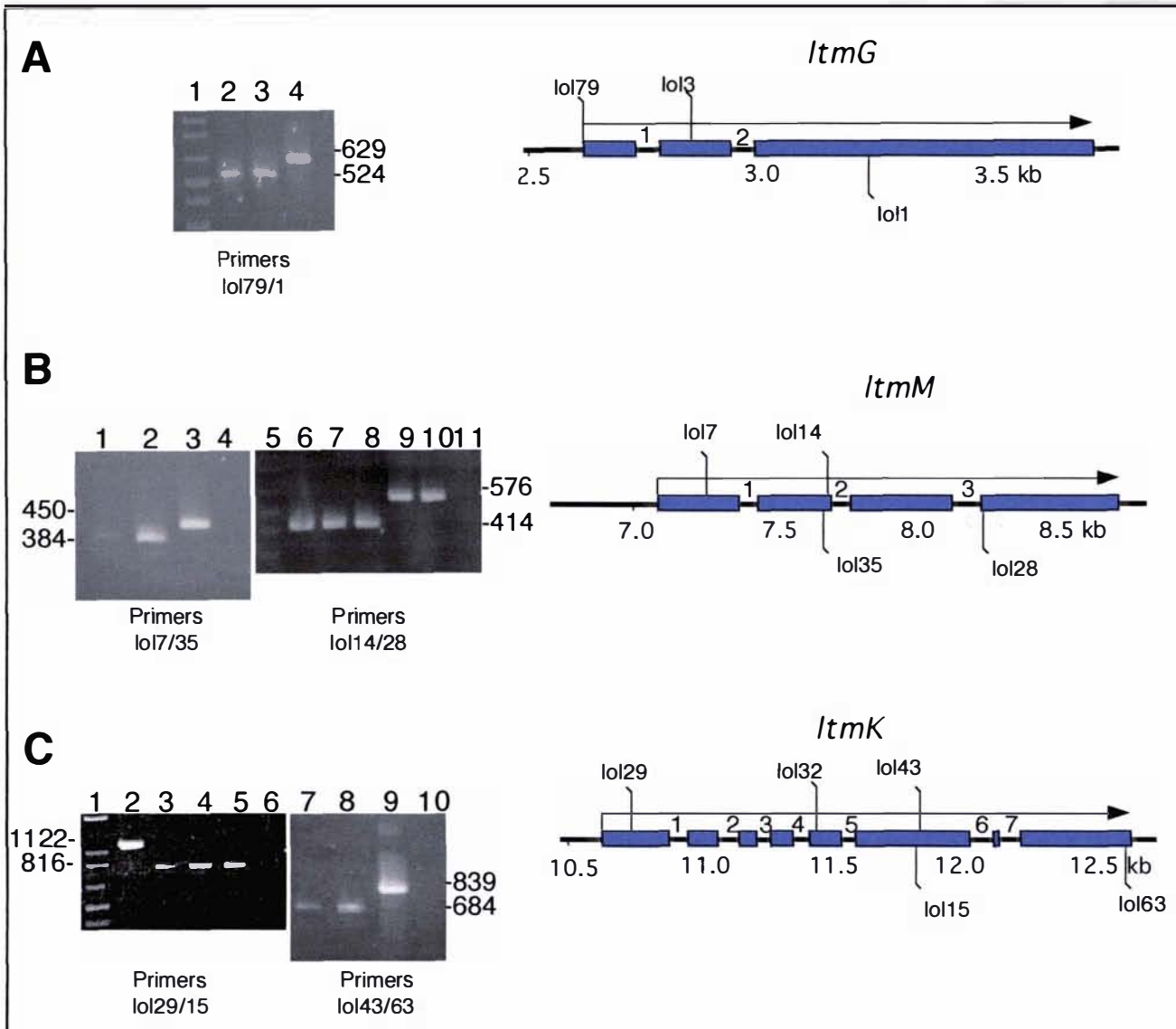


Figure 3.12 Gene structures of the *ltm* genes

Gene structures of the *ltm* genes. RT-PCR analysis was performed with RNA isolated from Lp19-infected perennial ryegrass. The cDNA was reverse transcribed from poly-A RNA using the Expand RT enzyme (Roche) (Section 2.6.2). All fragment sizes adjacent to gel photos are indicated in bp. The schematic diagrams of the genes have exons represented by blue boxes. The primers are positioned above or below the sequence for forward or reverse direction respectively. The approximate position of the gene within the cluster is indicated as kb. The introns are numbered between the exons. (**A**) The *ltmG* gene. Lane (1) 1kb+ ladder (Invitrogen); (2) 10^{-2} cDNA dilution; (3) 10^{-3} cDNA dilution; (4) Lp19 genomic DNA. (**B**) The *ltmM* gene. Lane (1 and 7) 10^{-2} cDNA dilution; (2 and 8) 10^{-3} cDNA dilution; (3 , 9 and 10) Lp19 genomic DNA; (4 and 11) water; (5) 1kb+ ladder (Invitrogen); (6) 10^{-1} cDNA dilution. (**C**) The *ltmK* gene. Lane (1) 1 kb+ ladder (Invitrogen); (2 and 9) Lp19 genomic DNA; (3, 4 and 5) 10^{-1} cDNA dilution; (6 and 10) water; (7) 10^{-2} cDNA dilution; (8) 10^{-3} cDNA dilution.

The following PCR reaction conditions were used: 5 ng genomic DNA or 5 µL of diluted cDNA, 1x *Taq* polymerase buffer (Roche), 50 µM each dNTP, 200 nM each primer, 0.5 U *Taq* polymerase in a reaction volume of 25 µL. The PCR amplification conditions for primer pair lol79/1, lol7/35, lol14/28 and lol43/63 were as follows: 94°C for 2 min; followed by 40 cycles of 94°C for 15 s, 60°C for 20 s, 72°C for 45 s; then one cycle of 72°C for 5 min. The PCR amplification conditions for primer combination lol29/15 were as follows: 94°C for 2 min; followed by 35 cycles of 94°C for 15 s, 60°C for 30 s, 72°C for 1 min 15 s; then one cycle of 72°C for 5 min.

Table 3.3 The *ltm* genes, intron analysis and comparisons to database sequences

Gene	Function	size (aa)	kDa	Intron				Top Database hit	Species	E value
				No.	phase	size	5'...3' splice sites			
ltmG	Geranylgeranyl diphosphate synthase	334	38	1	0	51	GTATTT...CAG	EAA72205	*F. graminearum*	6e-96
				2	1	54	GTAAGT...TAG			
ltmM	FAD dependent monooxygenase	472	53	1	1	66	GTAATA...TAG	*paxM*	*P. paxilli*	6e-96
				2	2	56	GTAGGT...TAG			
				3	0	106	GTATGT...CAG			
ltmK	P450 monooxygenase	533	61	1	0	66	GTATGT...CAG	EAA64305	*A. nidulans*	5e-84
				2	0	70	GTGAGG...TAG			
				3	0	52	GTATGT...CAG			
				4	2	67	GTATGC...TAG			
				5	2	51	GTACGT...CAG			
				6	1	81	GTTAGT...AAG			
				7	2	74	GTATGT...CAG			

(Fig. 3.12B). The resulting RT-PCR products were purified and sequenced with gene specific primers. The DNA sequence of the RT-PCR products was compared to the genomic sequence for confirmation of the intron splice sites (Table 3.3).

Of the three *ltmM* introns, the first two are conserved with those of *paxM* from *P. paxilli* (Young, et al., 2001), whereas all three introns are conserved with the *atmM* gene from *A. flavus* (Appendix 5.1). The top database BLASTP match is to PaxM with an E-value of 6e-96 where the overall sequence identity of LtmM to PaxM is 41.0% (across 456 amino acids). However, *A. flavus* AtmM is 42.2% identical (across 448 amino acids) to LtmM.

3.4.2 *ltmK*, a P450 monooxygenase

The sequence analysis of LtmK indicates it is more similar to PaxP than PaxQ with 31.3% (across 499 amino acids) and 23.4% (across 512 amino acids) identity at the amino acid level respectively. The BLASTP match of LtmK to PaxP is not as high as that seen when comparing LtmM to PaxM (E-value of 6e-96) indicating that *ltmK* is more likely to be a homologue of *paxP* than an orthologue. LtmK is likely to have a role in lolitrem biosynthesis as at least four cytochrome P450 steps have been predicted for lolitrem biosynthesis (Parker, Scott, 2004). The best database BLASTP match is to an *A. nidulans* sequence (accession number EAA64305) with an E-value of 5e-84.

Sequence analysis of *ltmK* predicts the presence of seven introns. These seven introns were confirmed by RT-PCR analysis using RNA isolated from endophyte-infected perennial ryegrass (Fig. 3.12C). The *ltmK* coding region was amplified from cDNA generated by random priming of mRNA using the primer combinations of lol29 and lol15 that would distinguish introns one through five, and primers lol43 and lol63 for introns six and seven (Fig. 3.12C). The resulting RT-PCR products were purified and sequenced with gene-specific primers then compared to the Lp19 genomic sequence for confirmation of the intron splice sites (Table 3.3). The *ltmK* gene has four introns, 1, 2, 4 and 7, that are conserved with placement and phase with the *paxP* introns 1, 2, 4 and 5, and three introns, 1, 4, and 7 that are conserved with placement and phase with the *paxQ* introns 2, 4, and 8 (Appendix 5.1).

3.5 Expression analysis of the *ltm* genes

The expression patterns of the *ltm* genes were determined by RT-PCR analysis and further supported by northern analysis (Section 2.6). RNA was isolated from Lp19 mycelia grown in PD liquid medium (Section 2.2.5) for approximately six days. A one step RT-PCR reaction (Invitrogen) with total RNA showed the three *ltm* genes were not expressed in culture (Fig. 3.13A). The expression of *tub2* was used as a control. These data are supported with the northern analysis, as shown in Figure 3.13C, where the *ltm* transcripts are not detected in culture.

The expression of the *ltm* genes *in planta* was determined by RT-PCR and Northern analysis (Fig. 3.13). RT-PCR was performed on mRNA isolated from Lp19-infected Nui perennial ryegrass using a random primed cDNA pool (Fig. 3.13B). Each *ltm* gene was amplified and compared to the amplification of the endophyte *tub2* and the perennial ryegrass β-tubulin genes. In comparison to the endophyte *tub2* transcript, which was just detected, the three *ltm* transcripts appear to be highly up-regulated *in planta*. A similar result is shown by northern analysis as seen in Figure 3.13C. The high level of expression of the three *ltm* genes detected *in planta* is evident especially when considering the endophyte is less than 1% of the plant material.

The DNA biomass of Lp19 in endophyte-infected perennial ryegrass was estimated to gauge the level of gene expression *in planta*. DNA was isolated from the pseudostems of Lp19-infected Nui perennial ryegrass, and the concentration of the endophyte was compared to DNA standards of Lp19 genomic DNA using Southern hybridisation with an *ltmK* fragment (Fig. 3.14). The hybridisation of the *ltmK* probe results in a single band of ~8.6 kb in both the Lp19 and Lp19-infected Nui perennial ryegrass genomic DNA. The intensity of the hybridising bands from the dilution series of the DNA from Lp19 mycelia was compared to those of the DNA from Lp19-infected Nui by densitometry with the Alpha Innotech Gel Documentation system. To be in the linear range of the X-ray film, only the 0.1 μg and 0.2 μg Lp19 standards, and the 5 μg and 10 μg Lp19-infected Nui samples were used in the analysis, as these DNA concentrations conformed to the expected ratio of ~1:2 when compared to each other. The hybridisation signal of each band was converted into a peak area and the ratio of the endophyte-alone to the endophyte-infected ryegrass peaks was determined. This ratio was then multiplied by the amount of each sample loaded and the reciprocal of the

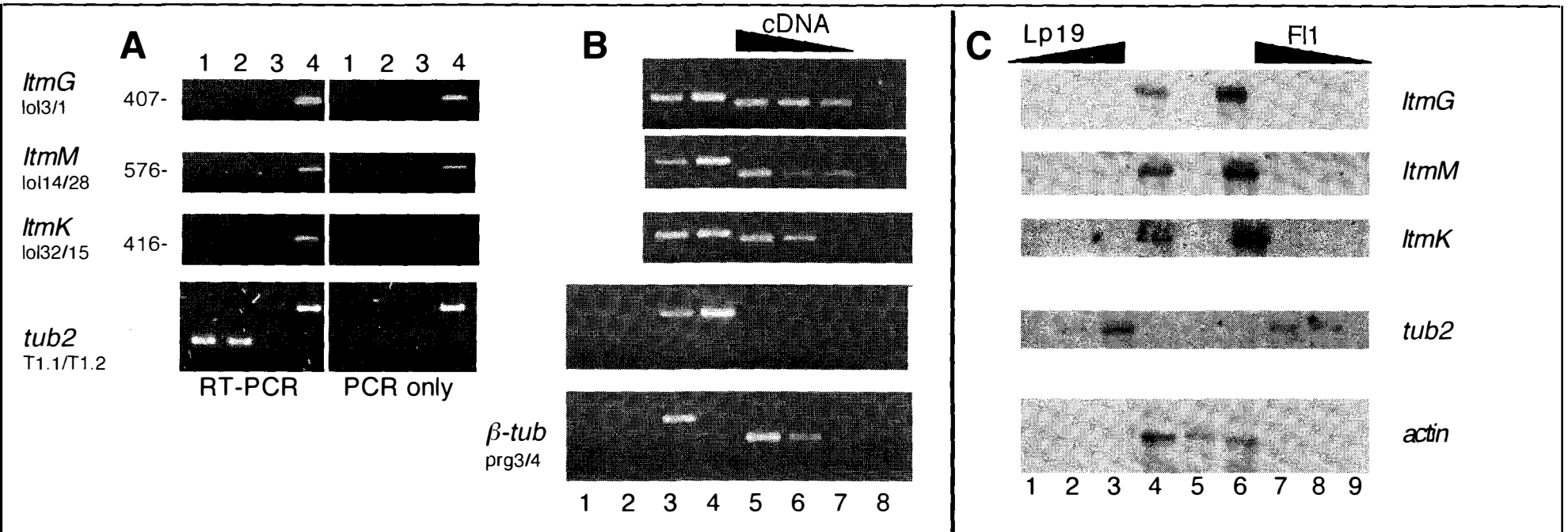
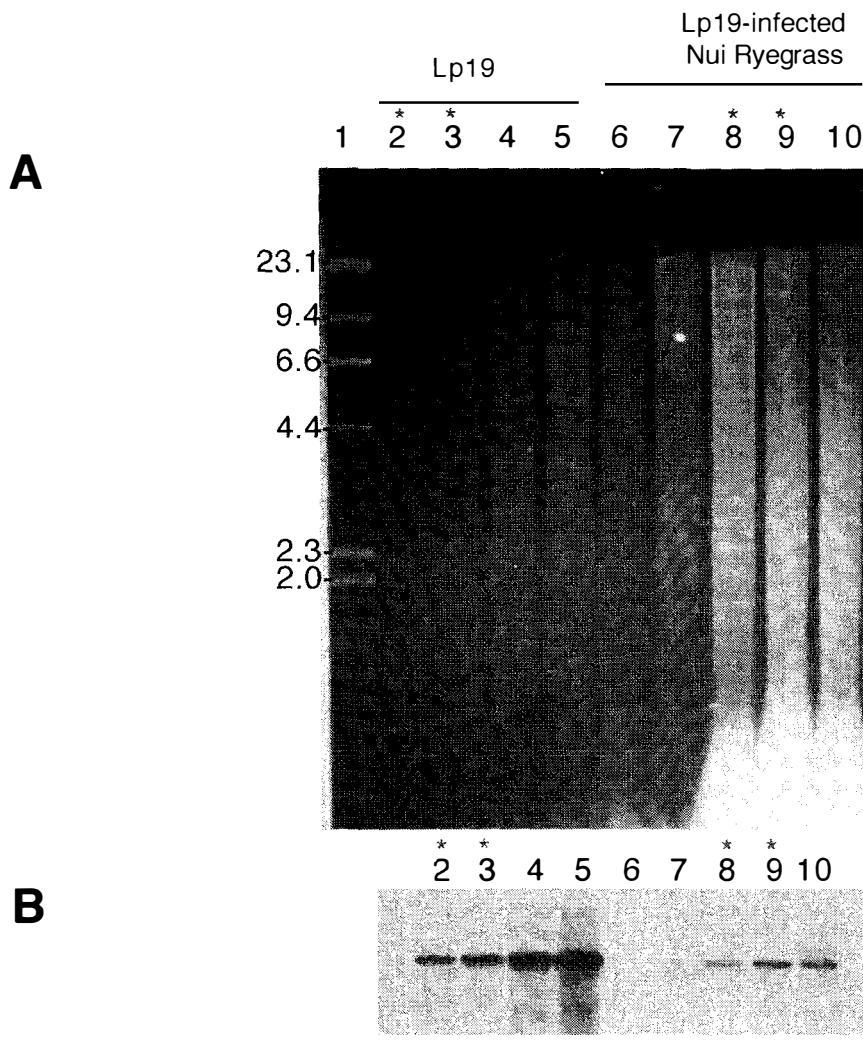


Figure 3.13 Expression analysis of the *ltm* genes

Expression analysis of the endophyte *ltmG*, *ltmM*, *ltmK* and *tub2* genes and the plant genes, *β-tubulin* and *actin*. Primer positions of the *ltm* genes are shown on Fig. 3.12. (**A**) RT-PCR analysis with RNA isolated from Lp19 mycelia (approximately 14 day old mycelia) using superscript one-step system (Invitrogen). Lanes (1) 1 μg Lp19 RNA; (2) 0.1 μg Lp19 RNA; (3) negative control; (4) Lp19 genomic DNA . The PCR conditions are identical to those shown in Fig. 3.10.

(**B**) RT-PCR analysis with RNA isolated from Lp19-infected Nui perennial ryegrass. The cDNA was made with mRNA isolated from Lp19-infected Nui perennial ryegrass. Lane (1) No RT enzyme control; (2) no mRNA control; (3) genomic DNA from Lp19-infected Nui perennial ryegrass; (4) Lp19 genomic DNA; (5) cDNA from Lp19-infected Nui perennial ryegrass, 10^{-1} dilution; (6) cDNA from Lp19 infected Nui ryegrass, 10^{-2} dilution; (7) cDNA from Lp19-infected Nui perennial ryegrass, 10^{-3} dilution; (8) negative control. The PCR reaction conditions are identical to those described in Fig. 3.12, with the following amplification conditions; 94°C for 2 min; followed by 35 cycles of 94°C for 15 s, 60°C for 30 s, 72°C for 45 s; then one cycle of 72°C for 10 min.

(**C**) Autoradiographs of Northern analysis with RNA isolated from Lp19 and Fl1 mycelia and Lp19- and Fl1-infected Nui perennial ryegrass. Lane (1) 0.15 μg Lp19 RNA; (2) 1.5 μg Lp19 RNA; (3) 15 μg Lp19 RNA; (4) 15 μg RNA from Lp19-infected Nui perennial ryegrass; (5) 15 μg RNA from endophyte-free Nui perennial ryegrass; (6) 15 μg RNA from Fl1-infected Nui perennial ryegrass; (7) 15 μg Fl1 RNA; (8) 1.5 μg Fl1 RNA; (9) 0.15 μg Fl1 RNA. The northerns were hybridised with the following ^{32}P-labelled fragments: *ltmG*, a cDNA product amplified with primers lol79 and lol1; *ltmM*, a cDNA product amplified with primers lol7 and lol28; *ltmP*, a cDNA product amplified with primers lol32 and lol8; *tub2*, a Lp19 genomic fragment amplified with primers T1.1 and T1.2; *actin*, a Nui genomic fragment amplified with primers actinF and actinR. The cDNA probes were cloned into pGEM-Teasy, digested with *Eco*RI, and the cDNA fragments were extracted from an agarose gel.



C

Lane	DNA sample	Conc. loaded	Scanned peak area	Ratio
2	Lp19	0.1 μg	805	
3	Lp19	0.2 μg	1484	1:1.8
8	Lp19/Nui	5 μg	300	
9	Lp19/Nui	10 μg	629	1:2.1

Figure 3.14 Endophyte biomass *in planta*

Southern analysis to estimate the endophyte biomass *in planta*. (**A**) Southern blot with Lp19 genomic DNA (lanes 2-5), and genomic DNA from Lp19-infected Nui ryegrass (lanes 6-10) digested with *Eco*RI. Each lane was loaded with the following amounts of DNA; (2) 0.1 μg Lp19; (3) 0.2 μg Lp19; (4) 0.5 μg Lp19; (5) 1.0 μg Lp19; (6) 1 μg Lp19/Nui; (7) 2 μg Lp19/Nui; (8) 5 μg Lp19/Nui; (9) 10 μg Lp19/Nui; (10) 20 μg Lp19/Nui. Lane (1) λ*Hin*dIII ladder. (**B**) Autogradiograph of the Southern blot from (**A**) hybridised with a ^{32}P-labelled *ltmK* probe amplified with lol32 and lol15. (**C**) A table showing the band densitometry as determined with an Alpha Innotech Gel Documentation system. The intensity of each hybridising band was converted into a peak area. The data from lanes with a * (lanes 2, 3, 8 and 9) were used as they are within the linear range and give comparable results between each like sample. Table 3.4 shows the loading concentrations, peak areas and the percentage endophyte *in planta*.

resulting number was converted into a percentage (Table 3.4). The average of the four comparisons (Table 3.4) provides an estimate of the Lp19 DNA biomass in endophyte-infected ryegrass of ~0.8%. This value of 0.8% endophyte is approximate, as variations in sampling will occur due to the condition of the grass and the amount of leaf blade removed from the pseudostem. However, the value is consistent with those seen in Section 3.6.2 and by Panaccione, et al., (2001).

3.6 Functional analysis of *ltmM*

Although the *ltm* genes were expressed *in planta* and correlate with strains that produce lolitrems, functional analysis was required to confirm the gene products had a role in lolitrem biosynthesis. A decision was made to disrupt the *ltmM* gene in the *E. festucae* strain Fl1 as this strain haploid, faster growing and has a better plant infection rate than Lp19. The *ltmM* gene was chosen for ease of disruption due to a large amount of available flanking sequence and because the disruption of the *P. paxilli* homologue, *paxM,* completely abolishes indole diterpene biosynthesis.

The *E. festucae* isolate Fl1 was screened for both hygromycin and geneticin sensitivity as these were the markers used for deletion and complementation of *ltmM,* respectively (Fig. 3.15). The growth rate of Fl1 on a range of hygromycin and geneticin concentrations was compared to that of Lp19 and E8 (Fig 3.15). Antibiotic concentrations were selected where minimal Fl1 growth was visible (Fig. 3.15). Based on these results, concentrations of 150 μg/mL of hygromycin and 200 μg/mL of geneticin were chosen for the deletion and complementation studies respectively.

The *E. festucae* Fl1 genomic region, from the start of *ltmG* to the end of *ltmK*, was directly sequenced with amplification products from each gene and intergenic region. The DNA sequence of the Fl1 *ltm* cluster is 99.9% identical with the Lp19 sequence, indicating there is very little sequence divergence between these two strains. There is a single C→T base difference (base 399 from the *ltmG* start codon) between the *E. festucae ltmG* and the *N. lolii ltmG*. Comparison of the *E. festucae ltmM* sequence to *N. lolii ltmM* shows two base changes of A→G at base 92 and T→C at base 249 from the *ltmM* start codon. Only the first transition results in a residue change with a conservative replacement of methionine (in *N. lolii ltmM*) to valine (in *E. festucae*

Table 3.4 The endophyte biomass *in planta*

A endophyte peak area	B endophyte/Nui peak area	C µg Endophyte/Nui	D µg Endophyte	E (A/B)*(C/D)	F 1/E	Endophyte biomass % endophyte[1]
805	300	5	0.1	134.17	0.007	0.75
805	629	10	0.1	127.98	0.008	0.78
1484	629	10	0.2	117.97	0.008	0.85
1484	300	5	0.2	123.67	0.008	0.81

[1]The following calculation, 1/[(A/B)(C/D)]100, was used to determine the endophyte biomass *in planta*

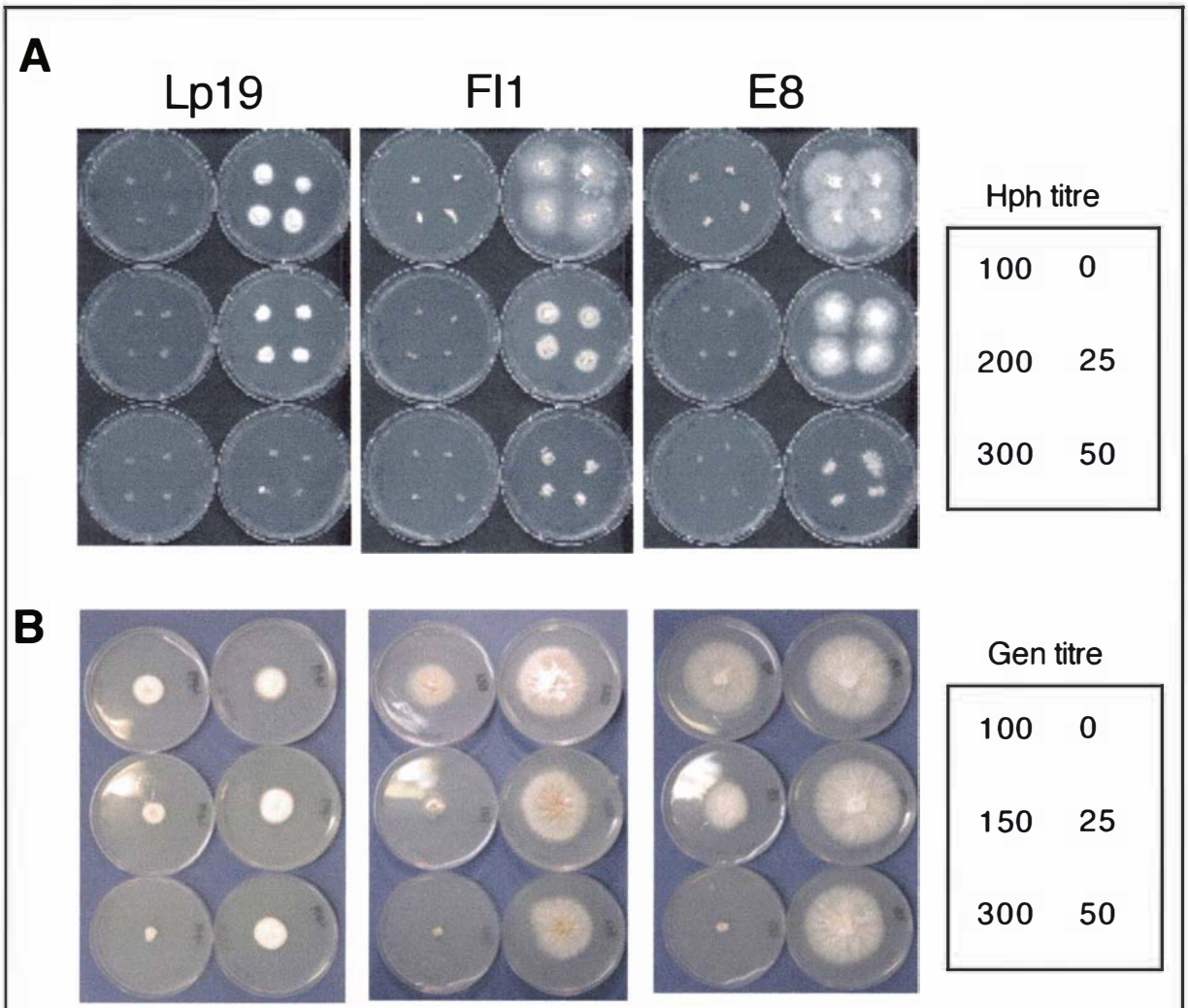


Figure 3.15 Determination of minimum inhibitory concentration of hygromycin and geneticin

Lp19, Fl1 and E8 growing on PD plates containing increasing concentrations of (**A**) hygromycin (Roche) or (**B**) geneticin (Invitrogen). The concentration of antibiotic in each plate is indicated in each panel key. All units of concentration are in μg/mL.

ltmM). The promoter region of the *N. lolii ltmM* and *E. festucae ltmM* has two differences, the first, a T→C change at base -356, is at a *Hin*dIII site that is absent from the Fl1 *ltmG-ltmM* intergenic region, and the second is at base -1038, where GAGA in Lp19 is expanded to GAGAGA in Fl1. The intergenic region between *ltmM* and *ltmK* has four base changes, while the *E. festucae ltmK* and *N. lolii ltmK* genes are 100% identical.

3.6.1 The *ltmM* deletion construct

The pCY39 plasmid made for deletion of the *E. festucae* Fl1 *ltmM* gene (Fig. 3.16) was constructed by sequentially ligating into pUC118 a 1.4 kb *Bam*HI fragment containing the hygromycin selectable marker (amplified with primers CY4 and CY5 from pCB1004; Carroll, et al., 1994), a 2.7 kb *Pst*I fragment (amplified with primers lol17 and lol18), followed by a 2.7 kb *Kpn*I fragment (amplified with primers lol48 and lol49). Each fragment required for the construct was amplified from Fl1 genomic DNA with the Expand High Fidelity PCR system (Roche) (Section 2.8) using the stated primers. The cloned orientation of each fragment was confirmed by sequence and restriction enzyme analysis. The pCY39 region required for deletion was linearised by amplification with the pUC forward and reverse primers using the Expand long template PCR system (Roche) (Section 2.8) and 5 μg of this product used to transform Fl1 (Section 2.9). After one month of growth the hygromycin-resistant transformants were colony-purified by three rounds of serial plating. DNA was extracted from each hygromycin resistant transformant (Section 2.3) and screened with primers lol148 and lol135 to confirm the integration of the plasmid (Fig. 3.17A). Transformants that contained an ectopic copy of the plasmid amplified both the 1.6-kb wild-type and the 1.4 kb integrating plasmid bands, whereas those that were homologous recombinants of the *ltmM* locus only contained the 1.4-kb integrated plasmid band. Transformants that were possible *ltm* deletions were further screened with the lol148 and lol135 primers, as well as with lol2 and lol34 (574 bp), lol7 and lol35 (448 bp) and lol147 and lol15 (317 bp), primers that amplify the left-hand side, *ltmM* gene and the right-hand side respectively (Fig. 3.17B). Any transformant with a deleted *ltmM* gene was screened further by Southern analysis to determine the pattern of integration of the pCY39 plasmid in the genome.

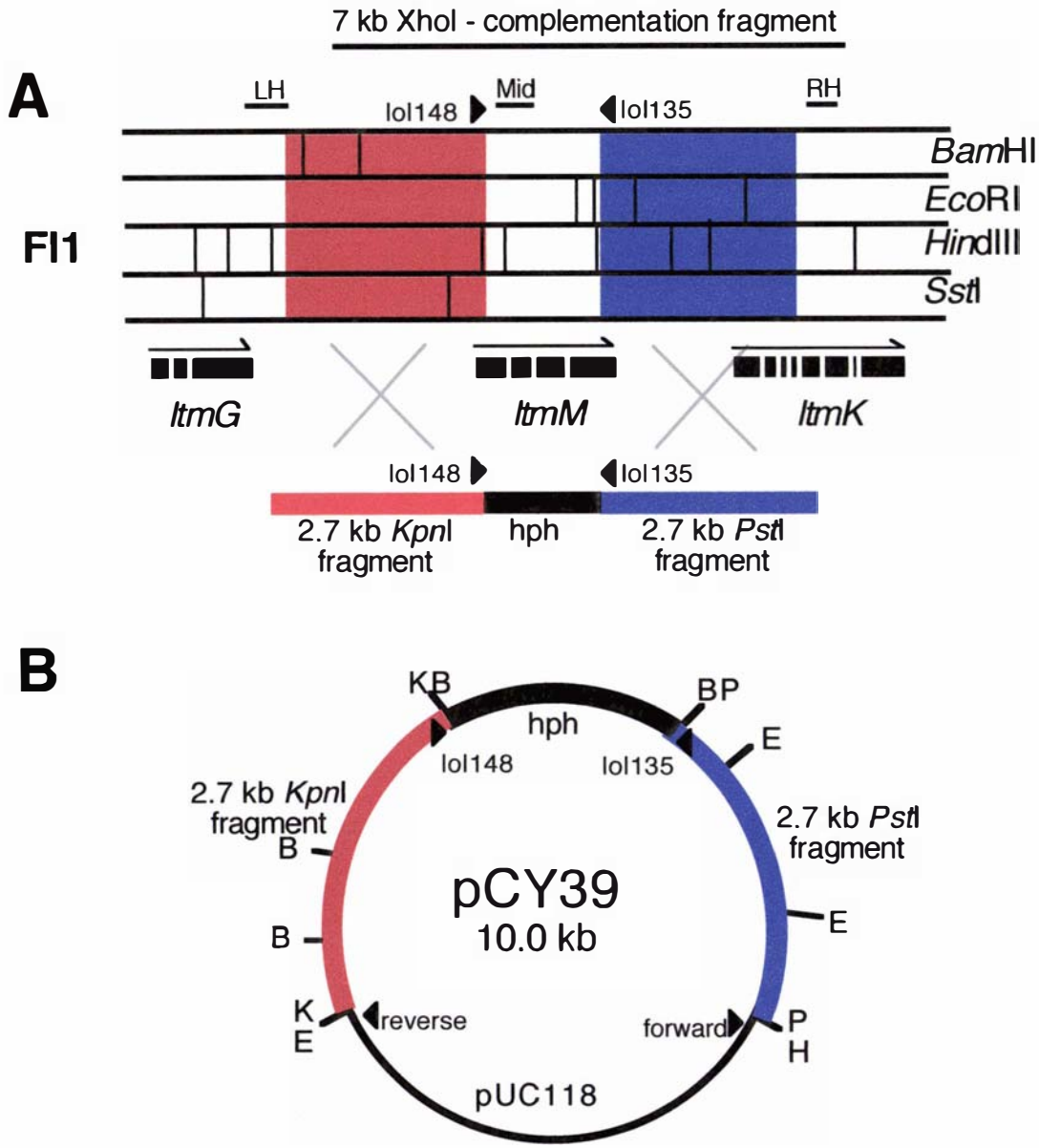


Figure 3.16 The ltmM-deletion construct

(**A**) A restriction enzyme map of the Fl1 *ltm* cluster. The red box is the 2.7 kb *Kpn*I fragment, the blue box is the 2.7 kb *Pst*I fragment. The LH, Mid and RH regions were amplified with primers lol2 and lol34, lol7 and lol35, and lol147 and lol15, respectively. Hph stands for the hygromycin resistance gene from pCB1004.

(**B**) A plasmid map of the deletion construct pCY39. The region targeted for deletion was amplified with the pUC forward and reverse primers using the Expand long template PCR system (Roche). Abbreviations for the restriction enzyme sites are: B = *Bam*HI, E = *Eco*RI, H = *Hin*dIII, K = *Kpn*I, P = *Pst*I. Not all restriction enzyme sites are shown in (**B**).

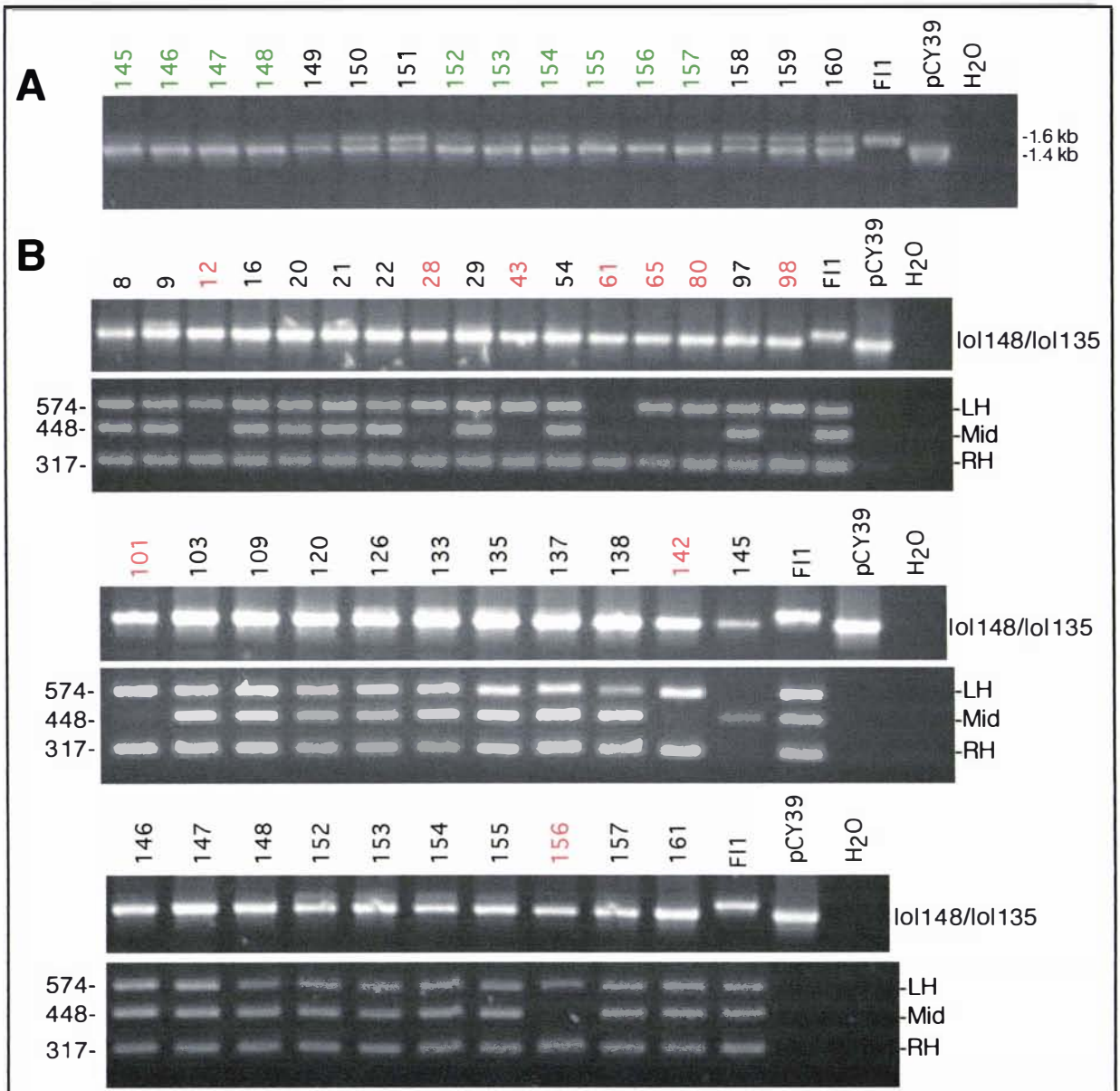


Figure 3.17 PCR screening for an *ltmM* deletion

PCR analysis of the *ltmM* Hyg^R transformants. (**A**) A selection of data from first round screening with primers lol148 and lol135. Transformants that gave a single 1.4 kb product, such as those defined with green numbers, were screened further as shown in (B). (**B**) All transformants that were possible *ltmM* deletions were rescreened with primer combination lol148 and lol135 and also screened with primer combinations to the left of the replacement construct (LH - primers lol2 and lol34), to the *ltmM* gene (Mid - primers lol7 and lol35) and to the right (RH - primers lol147 and lol15) of the replacement construct. See Figure 3.16 for primer locations. Any transformants that are labelled red are *ltmM* deletions and were screened further by Southern analysis. Fragment sizes are shown in bp. Transformant numbers have been abbreviated from CYFl1-M# to just the #.

Standard PCR conditions (Section 2.8) were used with the following PCR amplification conditions: For primer combination lol148 and lol135; 94^{o}C for 2 min, followed by 30 cycles of 94^{o}C for 15 s, 60^{o}C for 30 s, 72^{o}C for 2 min, then one cycle of 72^{o}C for 10 min. For the LH (lol2 and lol34), Mid (lol7 and lol35) and RH (lol147 and lol15) primer combinations, 94^{o}C for 2 min, followed by 30 cycles of 94^{o}C for 15 s, either 55^{o}C (for LH) or 60^{o}C (for Mid and RH) for 30 s, 72^{o}C for 45 s, then one cycle of 72^{o}C for 10 min. The LH, Mid and RH PCR amplifications were carried out separately and samples pooled for gel analysis.

A total of 159 arbitrarily selected hygromycin-resistant transformants were screened by PCR analysis. The PCR screening revealed 10 transformants, CYF11-M12, -M28, -M43, -M61, -M65, -M80, -M98, -M101, -M142 and -M156, that lacked the *ltmM* gene (Fig. 3.17B). During the screen for a homologous recombination event, one transformant, CYF11-M61, was identified that has a deletion of *ltmM* and the region to the left of *ltmG* (Fig. 3.17B). The extent of the CYF11-M61 deletion remains uncharacterised. However, based on hybridisation with a *ltmJ* fragment used in Section 3.9.5, the extent of the deletion is greater than 70 kb. One transformant, CYF11-M145, has an unusual PCR amplification pattern showing absence of both the left- and right-hand flanking regions (Fig. 3.17B). The integration pattern of CYF11-M145, detected by hybridisation with *hph* and *ltmM* fragments, is very complex (Fig. 3.18) as there are four *ltmM* hybridising bands and a strong hybridising fragment that represents a large tandem repeat. When screened by Southern analysis, five, CYF11-M28, -M43, -M65, -M142 and -M156, of the 10 transformants with *ltmM* deletions, were shown to have a single copy of the *hph* gene as indicated by hybridisation of a >23 kb band with a hph probe (Fig. 3.18). The frequency of replacement of *ltmM* was 6.2%, but only 3.1% of the total transformants were true single-copy replacements. Two independent deletion strains CYF11-M28 (PN2303) and CYF11-M142 (PN296), the *ltmMG* extended deletion mutant CYF11-M61 (PN2301), an ectopic mutant CYF11-M151 (PN2294), and wild-type F11 (PN2291) were used to artificially infect endophyte-free perennial ryegrass seedlings.

The same linear product amplified from pCY39 plasmid was used to transform Lp19 to delete the *NltmM* gene. An identical PCR screen to that described above was performed on 111 colony-purified hygromycin-resistant transformants but no Lp19-*ltmM* deletions were identified.

3.6.2 Alkaloid analysis of the *ltmM* knockouts

Perennial ryegrass seedlings were inoculated with the selected transformants by inserting mycelia into the meristematic region of an etiolated seedling (Section 2.10). Inoculated plants were grown for approximately six weeks and then screened for endophyte infection using an endophyte-specific polyclonal antibody (Section 2.10). Four to five endophyte-infected plants from each construct were maintained for lolitrem analysis. The rate of infection (Table 3.5), determined once the plants had established

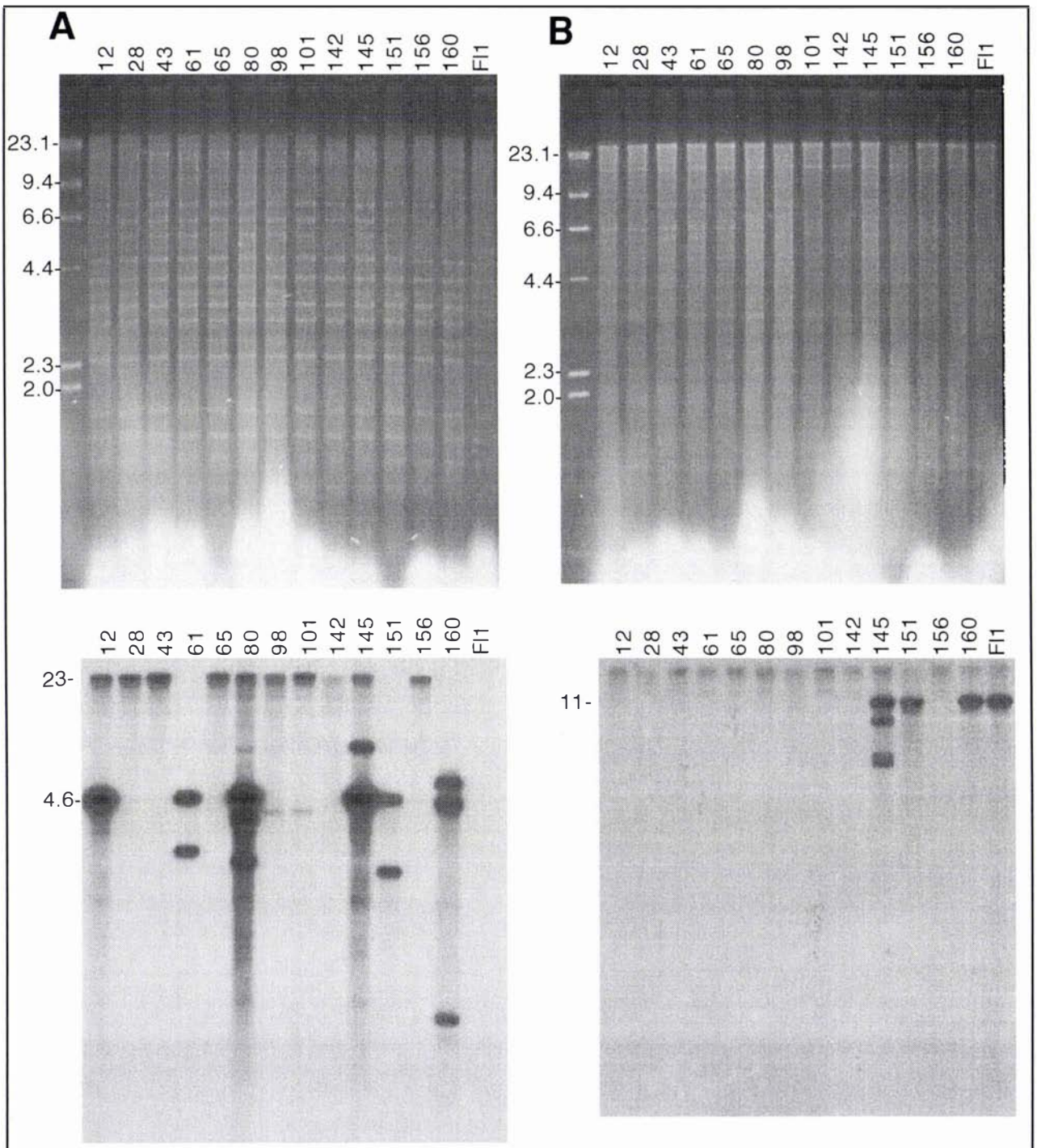


Figure 3.18 Southern analysis of the *ltmM*-deletion transformants

(**A**) *Eco*RI-digested genomic DNA from the CYFl1-M transformants hybridised with a ^{32}P-labelled fragment from the hygromycin resistance gene amplified with primers pUChph3 and pUChph4 from pCB1004. The presence of a tandem repeat of the integrating plasmid is seen as a 4.6 kb band. Integrations that were *ltmM* replacements have a >23 kb hybridising band.

(**B**) *Bam*HI-digested genomic DNA from the CYFl1-M transformants hybridised with a ^{32}P-labelled *ltmM* fragment amplified with primers lol7 and lol35. The wild-type band containing *ltmM* is ~11 kb. Transformants that lack the 11-kb band are *ltmM* deletions.

Transformant numbers have been abbreviated from CYFl1-M# to just the number. Ectopic integrants are CYLp19-M151 and -160.

Table 3.5 Rates of infection and alkaloid production of perennial ryegrass associations containing *ltmM* mutants

Strain	PN number	Fungal Type [1]	Number of plant/ associations	Infection Rate[2] (%)	Lolitrem[3] (ppm)	Ergovaline[3] (ppm)	Peramine[3] (ppm)
CYFII-M28	PN2303	Δ*ltmM*	5	20	0	0.4 - 1.3	30 - 40
CYFII-M61	PN2301	Δ*ltmMG*	4	17	0	0.7 - 3.3	24 - 41
CYFII-M142	PN2296	Δ*ltmM*	5	17	0	0.1 - 2.1	15 - 47
CYFII-M151	PN2294	Ectopic	5	17	4.4 - 16.7	0.5 - 1.2	21 - 55
FII	PN2291	wt	4	22	6.2 - 12.8	0.8 - 1.5	31 - 66
Endophyte Free		NA	3	NA	0	0	0

[1] Δ*ltmM* = *ltmM* insertional replacement, Δ*ltmMG* = extended deletion mutant, wt = Wild-type, NA = Not applicable.
[2] Infection rates were determined as a percentage of endophyte infected plants from the total surviving plants.
[3] Alkaloid analysis was performed mid-summer (January 2003).

reasonable growth, showed that each *ltmM*-deleted strain has a similar infection rate to that of wild-type Fl1. The phenotype of the association of the mutant endophyte with Nui perennial ryegrass was the same as wild-type Fl1 when plants were examined visually (Fig. 3.19) and by light microscopy (Fig. 3.20). The endophyte-infected plants were grown in a containment green house and were screened mid-summer (mid-January 2003) for the production of the alkaloids, lolitrem, ergovaline and peramine. All three alkaloids were extracted from pseudostems of endophyte-infected perennial ryegrass and analysed by HPLC. The alkaloid levels (Table 3.6 and summarised in Table 3.5) clearly showed that the endophyte strains with a deleted *ltmM* gene (CYFl1-M28, CYFl1-M61 and CYFl1-M142) are unable to produce lolitrem B (Fig. 3.21), but the levels of ergovaline and peramine were similar to Fl1 and the ectopic integrant CYFl1-M151 (Fig. 3.22, 3.23, Tables 3.5, 3.6). The *ltmM* deletion mutants, CYFl1-M28, CYFl1-M142 and CYFl1-M61, are also devoid of minor lolitrems A and E (Fig. 3.21). Variations of ergovaline and peramine levels between wild-type Fl1 and mutant associations were compared by statistical analysis using single factor ANOVA. The peramine levels between the wild-type and the mutant associations show no statistically significant variation (F factor 0.83; $P<0.52$). However, analysis of the first ergovaline sampling revealed a significant difference ($P<0.09$) between ergovaline levels in the CYFl1-M61-infected plants and the other associations. To ascertain whether this was a significant finding, the samples were rescreened for ergovaline and peramine late summer (end of February; 2003). The ergovaline levels of these samples, along with samples from a complementation experiment (Section 3.6.3), showed that the previous differences were not statistically significant.

The degree of endophyte colonisation *in planta* was determined, at the DNA level, by real-time PCR analysis of the endophyte *ggs1* gene and perennial ryegrass β-tubulin genes. DNA was isolated from wild-type Fl1 and pseudostems of endophyte-infected and uninfected perennial ryegrass. The endophyte and plant sample concentrations were determined by real-time PCR using a Roche Lightcycler (Section 2.8) and compared to the standard curves. To determine the concentration of the *ggs1* and β-tubulin (*tub*) genes, genomic DNA standards of endophyte and uninfected perennial ryegrass (perennial ryegrass plant G1138) were prepared by diluting boiled genomic DNA to concentrations of 1.5 ng/μL - 0.0015 ng/μL and 10 ng/μL - 0.01 ng/μL, respectively. The standards required for the endophyte *ggs1* amplification also contained perennial ryegrass genomic DNA at a concentration of 5 ng/μL. The *ggs1*

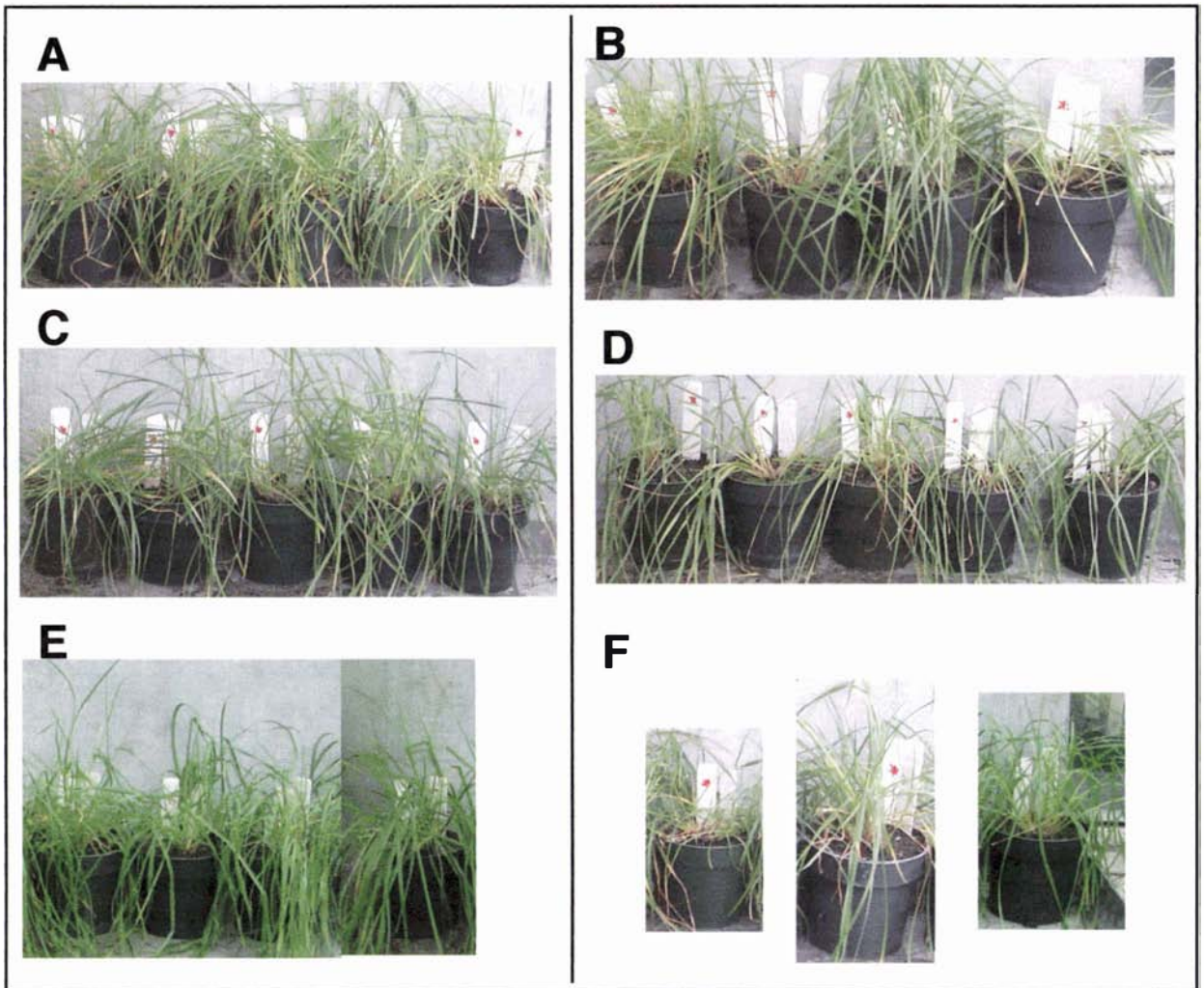


Figure 3.19 Visual inspection of the perennial ryegrass plants infected with FI1 carrying a deleted *ltmM* gene

Plant-endophyte associations. (A) CYFI1-M28-infected Nui perennial ryegrass, (B) CYFI1-M61-infected Nui perennial ryegrass, (C) CYFI1-M142-infected Nui perennial ryegrass, (D) CYFI1-M151-infected Nui perennial ryegrass, (E) FI1-infected Nui perennial ryegrass, (F) endophyte-free perennial ryegrass.

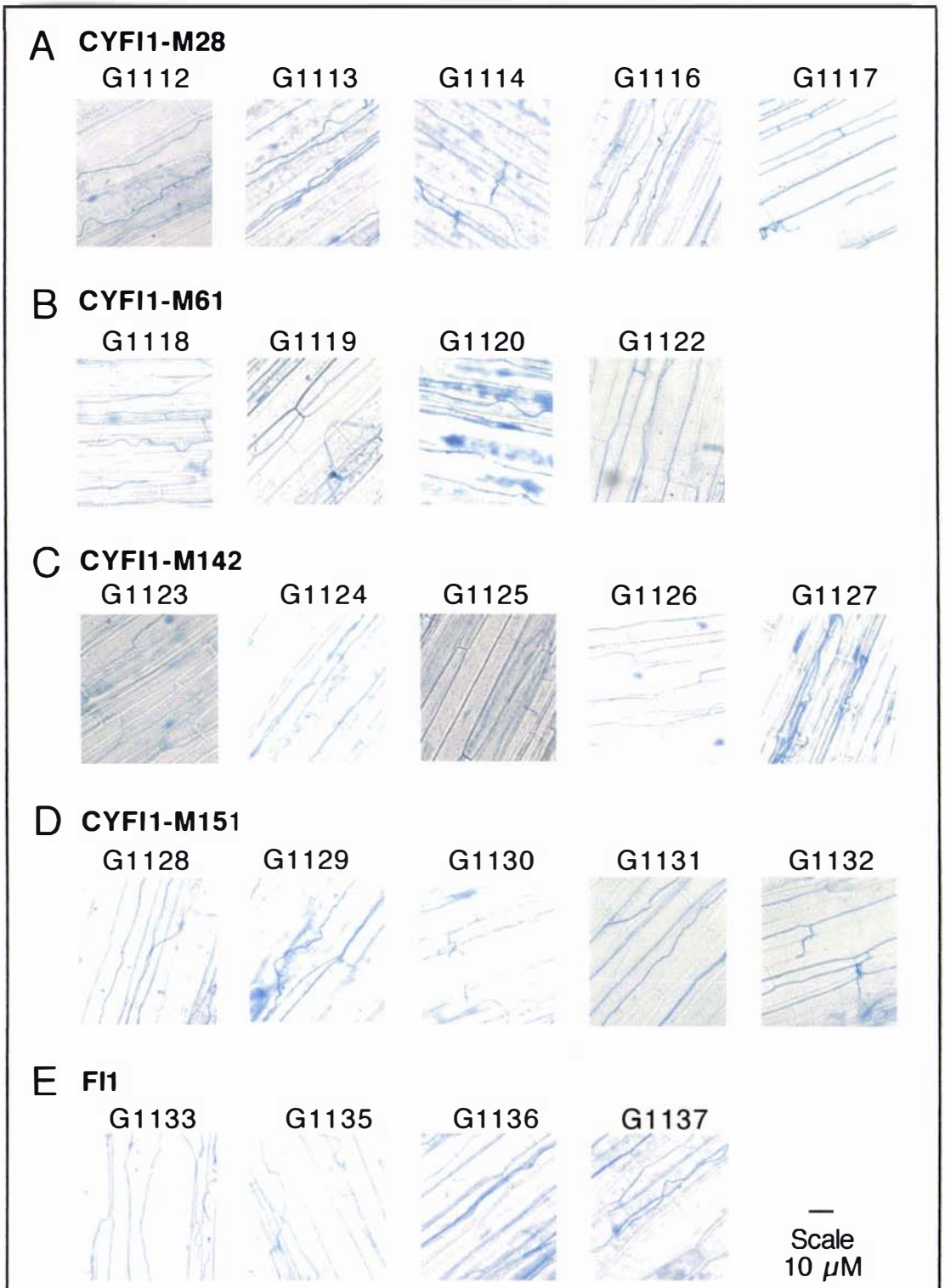


Figure 3.20 Light microscopy of perennial ryegrass leaves infected with FI1 *ItmM*-deletion mutants

Longitudinal sections of epidermal leaf peels of Nui perennial ryegrass infected with the *ItmM* deletions (**A**) CYFI1-M28; (**B**) CYFI1-M61; (**C**) CYFI1-M142; (**D**) ectopic CYFI1-M151; or (**E**) wildtype FI1 stained with aniline blue. Leaf peels of pseudostems were treated with the aniline blue stain (Section 2.10) and subjected to microscopy. The endophyte is seen as the dark blue stain between the plant cells.

Table 3.6 Alkaloid analysis of endophyte infected ryegrass

Plant association number	Endophyte strain	Genotype[1]	[2]Lolitrem (ppm)	[2]Ergovaline (ppm)	[2]Peramine (ppm)	[3]Ergovaline (ppm)	[3]Peramine (ppm)
G1112	CYFI1-M28	Δ*ltmM*	0	0.5	40	0.2	88
G1113	CYFI1-M28	Δ*ltmM*	0	0.7	30	0.2	85
G1114	CYFI1-M28	Δ*ltmM*	0	1.3	30	1.2	135
G1116	CYFI1-M28	Δ*ltmM*	0	0.7	33	1.0	122
G1117	CYFI1-M28	Δ*ltmM*	0	0.4	30	0.6	109
G1118	CYFI1-M61	Δ*ltmMG*	0	3.3	38	2.6	83
G1119	CYFI1-M61	Δ*ltmMG*	0	0.8	24	1.4	81
G1120	CYFI1-M61	Δ*ltmMG*	0	3.3	41	1.3	75
G1122	CYFI1-M61	Δ*ltmMG*	0	0.7	32	0.3	101
G1123	CYFI1-M142	Δ*ltmM*	0	0.3	15	0.3	45
G1124	CYFI1-M142	Δ*ltmM*	0	0.1	37	0.1	75
G1125	CYFI1-M142	Δ*ltmM*	0	0.5	29	0.1	77
G1126	CYFI1-M142	Δ*ltmM*	0	2.1	47	0.9	87
G1127	CYFI1-M142	Δ*ltmM*	0	0.6	29	0.2	95
G1128	CYFI1-M151	Ectopic	7.7	1.2	26	0.2	85
G1129	CYFI1-M151	Ectopic	4.4	0.7	21	0.5	67
G1130	CYFI1-M151	Ectopic	16.7	1.2	33	1.7	86
G1131	CYFI1-M151	Ectopic	8.5	0.5	55	0.6	96
G1132	CYFI1-M151	Ectopic	7.1	0.8	34	0.3	93
G1133	FI1	wt	8.4	1.5	38	1.2	77
G1135	FI1	wt	7.9	1.1	31	1.1	108
G1136	FI1	wt	12.8	0.8	40	2.0	118
G1137	FI1	wt	6.2	0.8	66	1.3	88
G1115	N/A	Endophyte free	0	0	0	nt	nt
G1121	N/A	Endophyte free	0	0	0	nt	nt
G1138	N/A	Endophyte free	0	0	0	nt	nt

[1] Δ*ltmM* = *ltmM* insertional replacement, Δ*ltmMG* = extended deletion mutant, wt = wild-type FI1
[2]Samples were tested mid January (2003). HPLC traces are shown in Figures 3.21-3.23.
[3]Samples were tested late February (2003)
nt = not tested

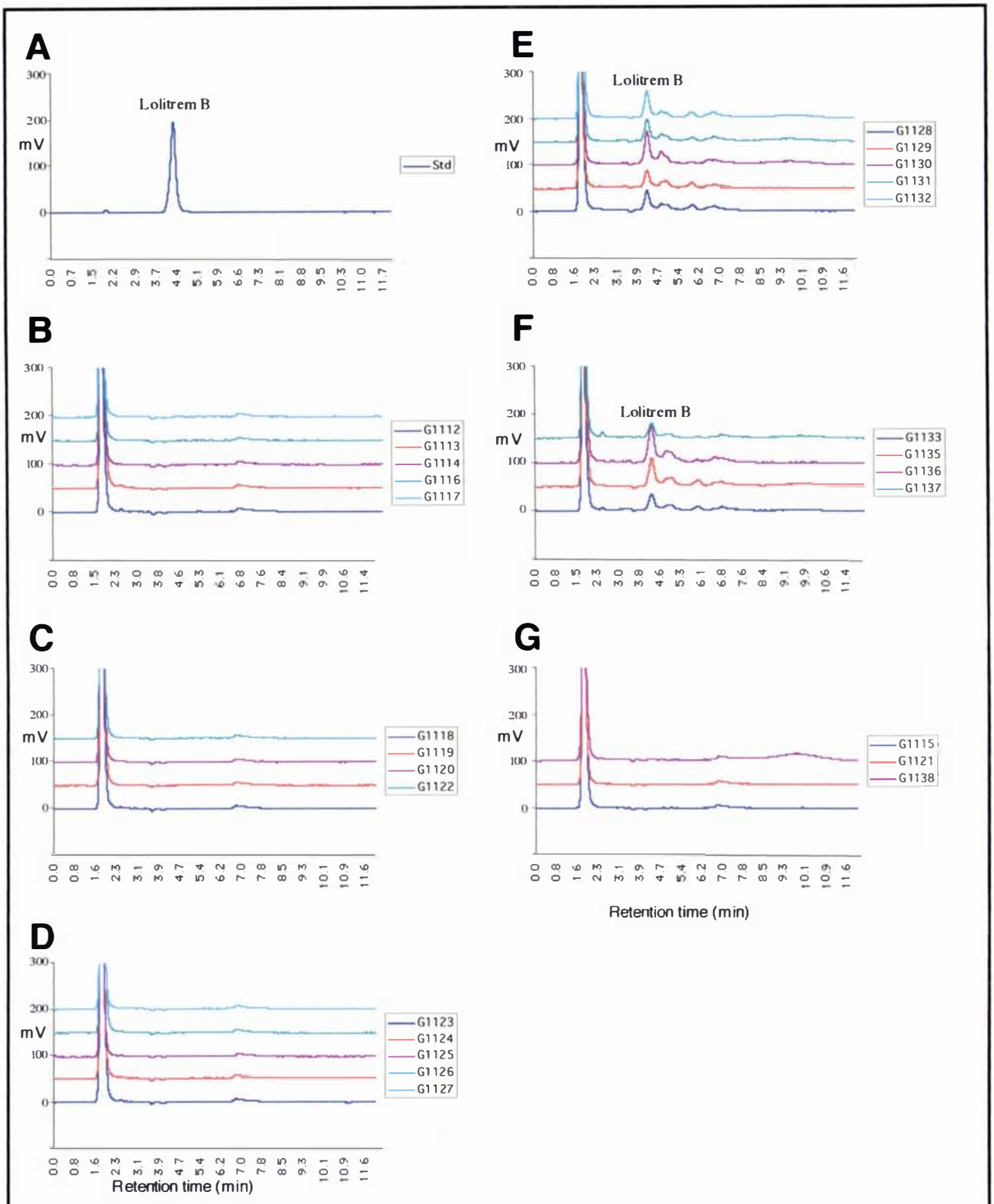


Figure 3.21 HPLC analysis for lolitrem B in the *ltmM*-deletion mutants

HPLC analysis of indole-diterpene extracts from pseudostems of endophyte-infected Nui perennial ryegrass. (**A**) Lolitrem B standard (8.4 ng). Samples of Nui perennial ryegrass infected with (**B**) CYFI1-M28, (**C**) CYFI1-M61, (**D**) CYFI1-M142, (**E**) CYFI1-M151, (**F**) FI1, (**G**) uninfected. A key to individual endophyte/ryegrass association numbers is given to the right of each graph. The peaks eluting later than lolitrem B are lolitrems A and E. Experimental details are given in Section 2.11.1.

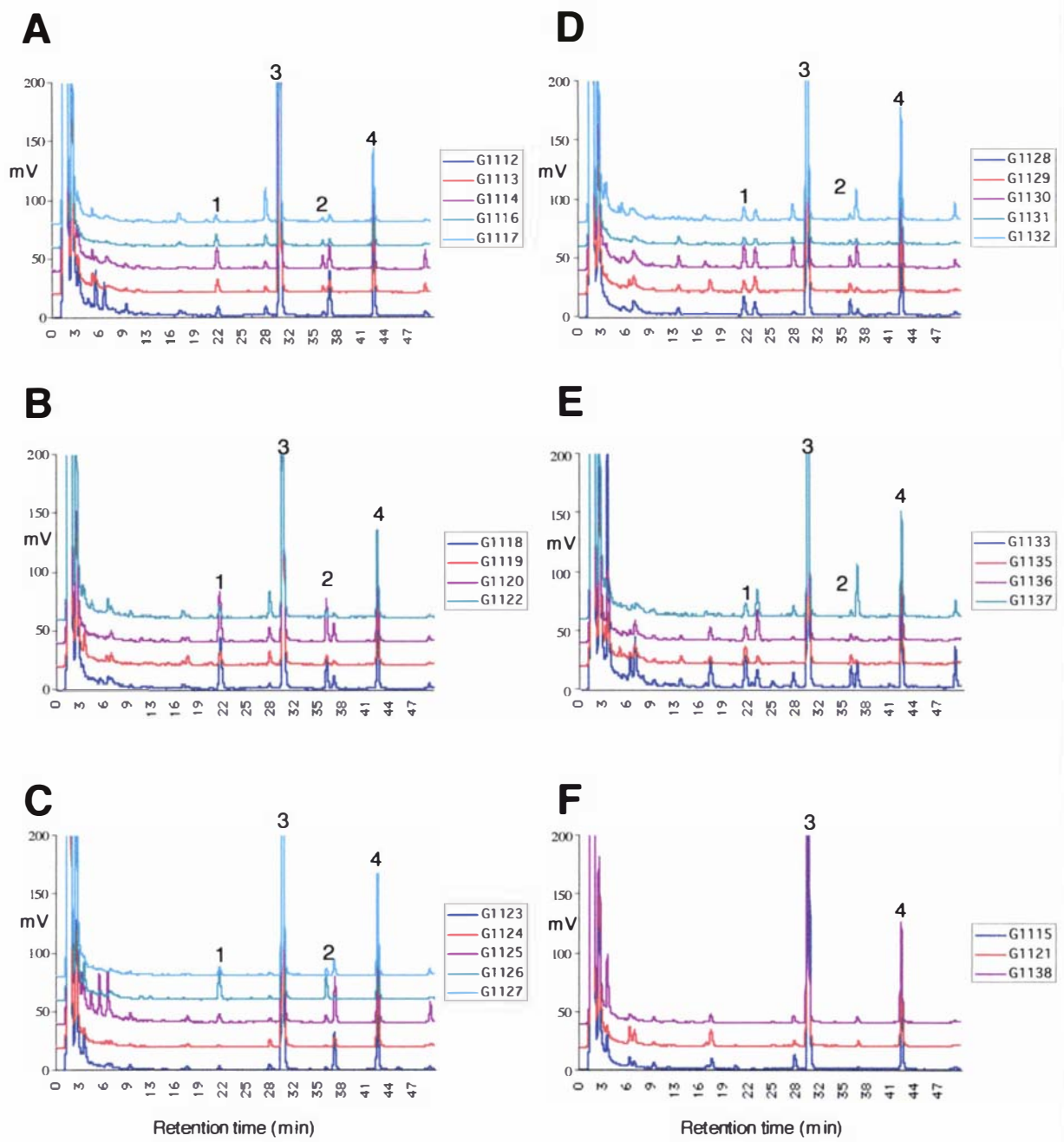


Figure 3.22 HPLC analysis for ergovaline in the *ltmM*-deletion mutants

HPLC analysis for the presence of ergovaline in pseudostems of endophyte infected Nui perennial ryegrass. Samples of Nui perennial ryegrass infected with (**A**) CYFl1-M28, (**B**) CYFl1-M61, (**C**) CYFl1-M142, (**D**) CYFl1-M151, (**E**) Fl1, and (**F**) endophyte free. Peaks labelled 1 and 2 are ergovaline and isomer ergovalinine respectively. Peaks labelled 3 and 4 are the internal standards ergotamine and ergotaminine that were added to the samples before extraction. A key to the individual endophyte/ryegrass association numbers is given to the right of each graph. Experimental details are given in Section 2.11.2 and 2.11.3.

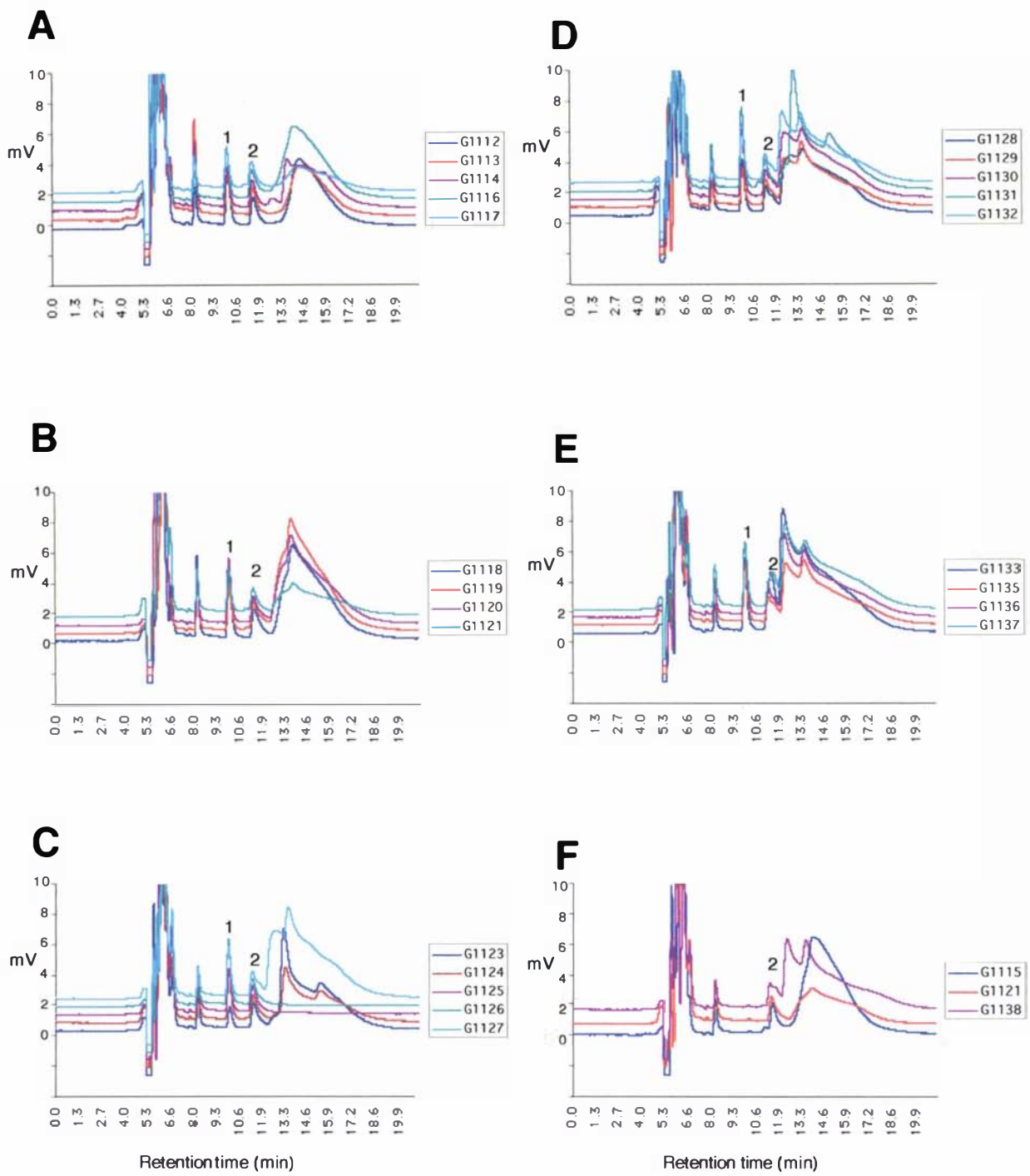


Figure 3.23 HPLC analysis for peramine in the *ltmM*-deletion mutants

HPLC analysis for the presence of peramine in pseudostems of endophyte infected Nui perennial ryegrass. Samples of Nui perennial ryegass infected with (**A**) CYFl1-M28, (**B**) CYFl1-M61, (**C**) CYFl1-M142, (**D**) CYFl1-M151, (**E**) Fl1 and (**F**) endophyte-free. The peak labelled 1 is peramine. The peak labelled 2 is the internal standard of homoperamine that was added to the samples before extraction. A key to the individual endophyte/ryegrass association numbers is to the right of each graph. Experimental details are given in section 2.11.2 and 2.11.4.

and *tub* genes were amplified with primers endo1 and endo2 (*ggs1*), or prg12 and prg13 (*tub*), using the diluted genomic DNA standards as templates. At the end of each PCR amplification cycle the level of fluorescence was detected and subsequently plotted against the cycle number (Fig. 3.24). A standard curve for the amplification of each gene was prepared by plotting the log concentration of the DNA standards against the cycle number cross-over point (Fig. 3.24).

To determine the DNA concentration of the endophyte and perennial ryegrass from endophyte/plant associations, genomic DNA was boiled, diluted to 5 ng/μL and amplified with primer pairs endo1 and endo2, and prg12 and prg13 (Fig. 3.24, 3.25). At the end of the PCR programme a melting curve was plotted for each sample to show that the PCR amplification was specific (Appendix 5.5.1). Perennial ryegrass samples that were endophyte-free showed very low levels of PCR amplification with the endophyte primers and, moreover, these samples did not have the same melting curve as the samples that contained a genuine *ggs1* amplification product (Appendix 5.5.2). The G1118 sample had a lower DNA concentration than the other samples with both the endophyte and the ryegrass primers. The endophyte DNA biomass was determined as a percentage of endophyte in the plant sample (Table 3.7). The endophyte concentrations of the mutant associations averaged about 1%, consistent with associations containing wild-type Fl1 and Lp19 (Section 3.5). Therefore, the absence of lolitrem B in the deletion strains is due to the effect of the deleted *ltmM* gene and not to a difference in endophyte biomass. This result is confirmed by the presence of normal levels of peramine and ergovaline produced by the mutant endophyte associations (Tables 3.5, 3.6). The deletion of the *ltmM* gene has no effect on the ability of the mutants to colonise the host.

3.6.3 Complementation of *ltmM*

Two constructs were made to test whether the *ltmM* deletions CYFl1-M28 and CYFl1-M61 could be complemented. A *ltmM* complementation construct, pCY40, was made by cloning a 7 kb *Xho*I fragment from λCY218 into the *Xho*I site of pII99 (Fig. 3.16A; 3.26A). The 7-kb *Xho*I fragment contains 2.2 kb of 5' and 3 kb of 3' sequence flanking the *ltmM* gene. A *paxM* complementation construct, pCY41, was made by cloning a 3.7 kb *Bgl*II fragment from λCY46 (Fig. 3.26B) into the *Bgl*II site of pII99. The pCY41 plasmid contains 1.5 kb 5' and 0.66 kb 3' sequence flanking the *paxM* gene. Protoplasts

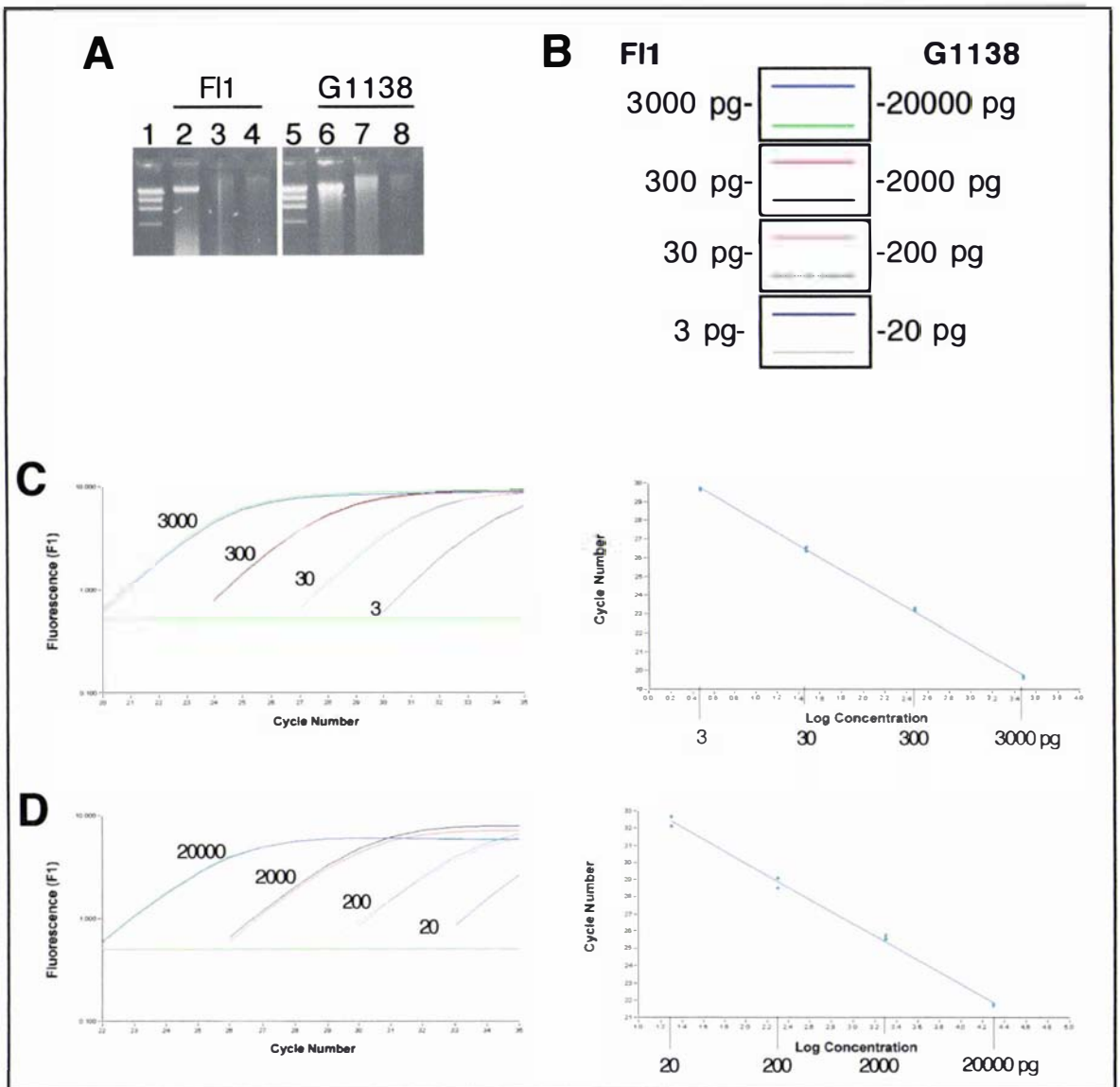


Figure 3.24 Standard curves for estimation of endophyte DNA biomass

(**A**) Genomic DNA from Fl1 and endophyte-free perennial ryegrass, G1138. Lane (1) λ*Hin*dIII standard, (2) Fl1 genomic DNA, (3) boiled Fl1 genomic DNA, (4) 90 ng of boiled Fl1 genomic DNA, (6) G1138 genomic DNA, (7) boiled G1138 genomic DNA, (8) 90 ng of boiled G1138 genomic DNA. The samples in lanes 2, 3, 6 and 7 were to determine the quality of the DNA and are of unknown concentrations. Lanes 4 and 8 have the same amount of DNA loaded, which confirmed that the fluorometer gave equivalent concentration readings for the boiled Fl1 and G1138 genomic DNA.

(**B**) The identification colours for the amplification products of the Fl1 and G1138 standards.

(**C**) Real-time PCR amplification of endophyte *ggs1* gene using the Fl1 DNA standards with primers endo1 and endo2. A graph plotting cycle number versus fluorescence and subsequent conversion of this into a standard curve plotting concentration of *ggs1* PCR product (amplified from the Fl1 DNA standards) versus cycle number. The concentrations are in pg.

(**D**) Real-time PCR amplification of Nui β-tubulin gene using the G1138 DNA standards with primers prg12 and prg13. A graph plotting cycle number versus fluorescence and subsequent conversion of this into a standard curve plotting concentration of the *tub* PCR product (amplified from the G1138 DNA standards) versus cycle number. The concentrations are in pg.

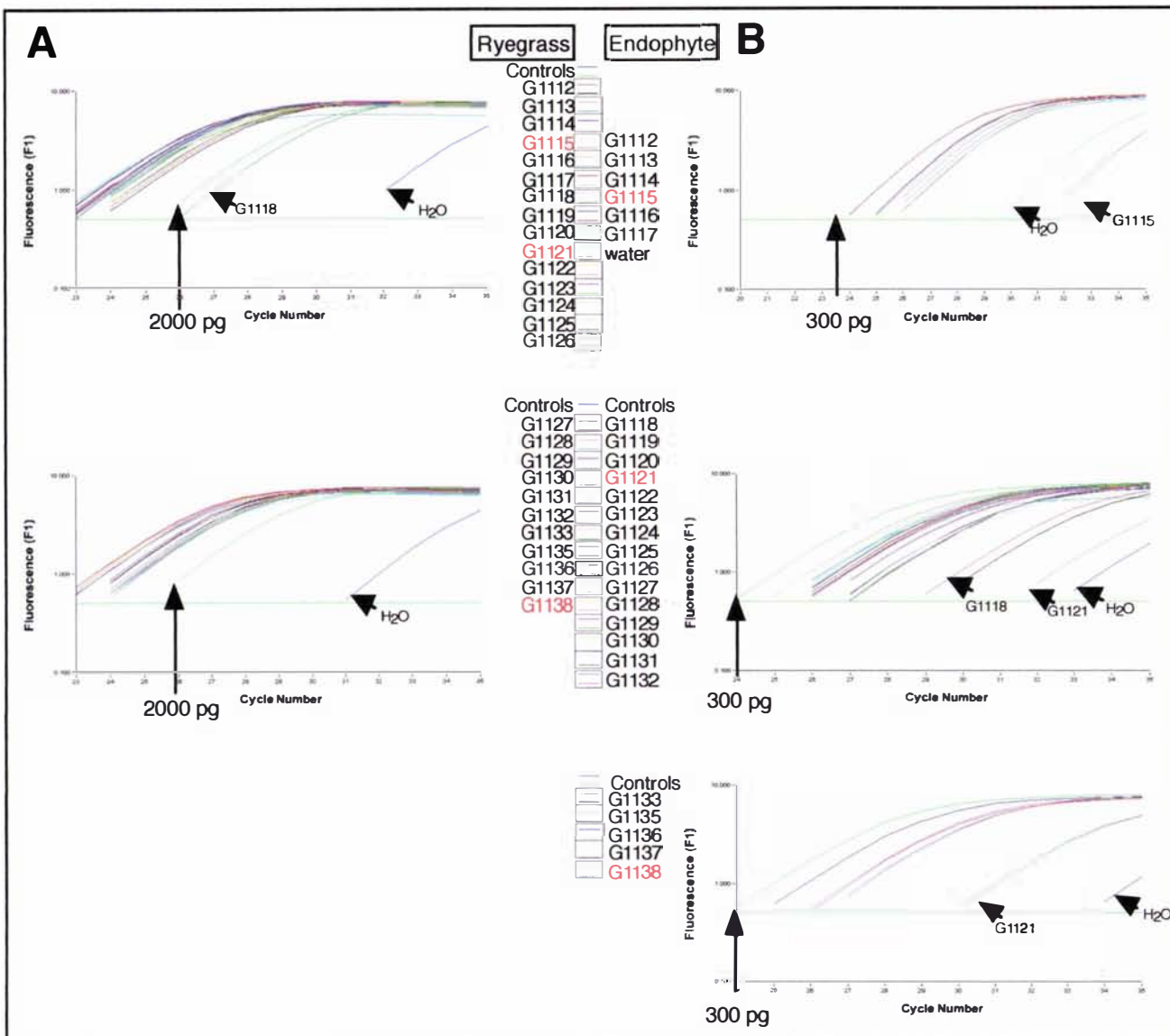


Figure 3.25 Real-time PCR analysis to determine endophyte DNA biomass

(**A**) Levels of plant β-tubulin gene detected by real-time PCR using 10 ng of boiled DNA isolated from endophyte-infected pseudostems and primers prg12 and prg13. All samples were compared to the ryegrass DNA standard curve (Fig. 3.24D) by calibration of the internal standard of 2000 pg G1138 DNA for each run.

(**B**) Levels of the endophyte *ggs1* detected by real-time PCR using 10 ng of boiled DNA isolated from endophyte-infected pseudostems and primers endo1 and endo2. All samples were compared to the endophyte DNA standard curve (Fig. 3.24D) by calibration of the internal standard of 300 pg Fl1 DNA for each run.

Samples were tested in duplicate. A key to the sample colours is shown between A and B. The outlier samples are indicated with arrows. The blue control for each sample set is water, the green control for each sample set is the G1138 2000 pg DNA standard or the Fl1 300 pg DNA standard for the perennial ryegrass or endophyte samples, respectively. The samples labelled red in the key are endophyte-free. The PCR amplification products for the endophyte-free samples, including the water control, were non-specific as determined by the sample melting curve (Appendix 5.5). The numbers generated from the real-time PCR have been converted into percentage endophyte as shown in Table 3.7.

Table 3.7 Endophyte biomass *in planta*

Plant association number	Endophyte strain	Genotype[3]	Endophyte average conc pg Endo1/2	Std deviation +/-	Plant average conc pg PRG12/13	Std deviation +/-	% endophyte[2]	Average associations
G1112	CYFI1-M28	Δ*ltmM*	58.7	4.7	6257	521	0.9	1.0
G1113	CYFI1-M28	Δ*ltmM*	78.6	5.7	10810	568	0.7	
G1114	CYFI1-M28	Δ*ltmM*	185.1	0.0	12800	382	1.4	
G1116	CYFI1-M28	Δ*ltmM*	91.8	3.3	7314	3	1.3	
G1117	CYFI1-M28	Δ*ltmM*	50.5	1.2	13070	329	0.4	
G1118	CYFI1-M61	Δ*ltmMG*	11.2	2.5	1494	222	0.7	0.8
G1119	CYFI1-M61	Δ*ltmMG*	47.8	0.9	10640	442	0.4	
G1120	CYFI1-M61	Δ*ltmMG*	43.8	4.2	10000	1206	0.4	
G1122	CYFI1-M61	Δ*ltmMG*	112.6	4.1	7770	215	1.4	
G1123	CYFI1-M142	Δ*ltmM*	69.2	1.7	9422	52	0.7	1.1
G1124	CYFI1-M142	Δ*ltmM*	42.0	1.1	9655	245	0.4	
G1125	CYFI1-M142	Δ*ltmM*	94.5	5.2	9158	382	1.0	
G1126	CYFI1-M142	Δ*ltmM*	154.4	18.5	12090	1843	1.3	
G1127	CYFI1-M142	Δ*ltmM*	128.0	6.6	6912	234	1.9	
G1128	CYFI1-M151	Ectopic	95.7	3.0	10330	292	0.9	1.1
G1129	CYFI1-M151	Ectopic	104.5	2.0	7924	445	1.3	
G1130	CYFI1-M151	Ectopic	154.7	5.3	10300	743	1.5	
G1131	CYFI1-M151	Ectopic	38.4	2.3	4701	37	0.8	
G1132	CYFI1-M151	Ectopic	86.0	1.1	11670	833	0.7	
G1133	FI1	wt	79.8	3.7	5463	674	1.5	1.3
G1135	FI1	wt	51.1	2.5	5149	251	1.0	
G1136	FI1	wt	177.4	0.0	10370	693	1.7	
G1137	FI1	wt	68.2	6.1	6313	135	1.1	
G1115[1]	NA	Endophyte free	0.9	0.4	9392	1721	0.0	0.0
G1121[1]	NA	Endophyte free	3.2	1.1	15150	765	0.0	
G1138[1]	NA	Endophyte free	4.4	0.3	11950	412	0.0	

[1]The endophyte amplification product was nonspecific and therefore the % endopyhte associations are 0.

[2]There is no statistically significant variation (P<0.05) between sample groups as determined by single factor ANOVA.

[3]Δ*ltmM* = *ltmM* insertional replacement, Δ*ltmMG* = extended deletion mutant, wt = wild-type FI1, NA not applicable

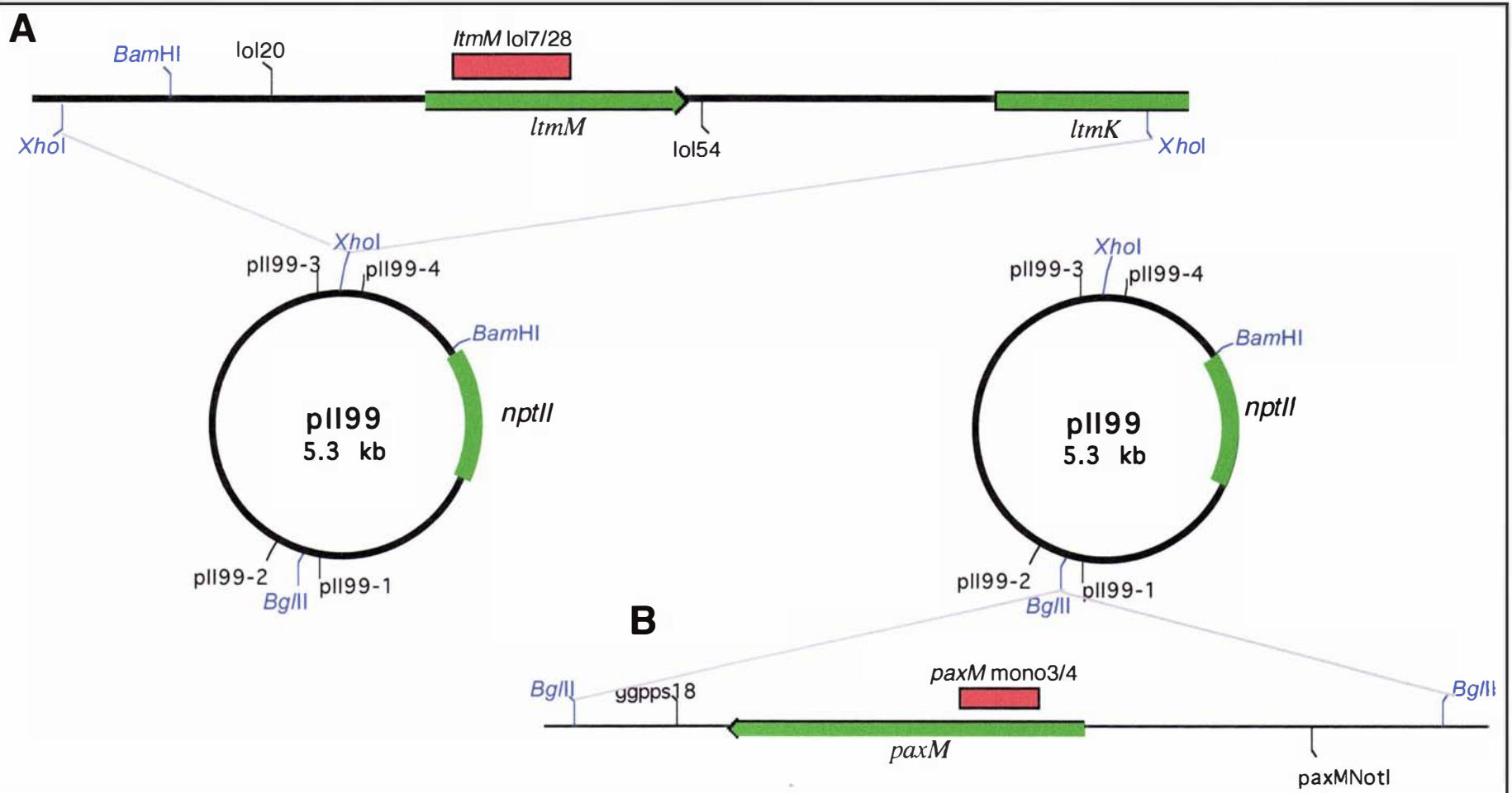


Figure 3.26 *ltmM* and *paxM* complementation constructs

Constructs used for *ltmM* and *paxM* complementation analysis. **(A)** pCY40 contains the 7 kb *Xho*I fragment from λCY218 cloned into the *Xho*I site of pll99. **(B)** pCY41 contains the 3.7 kb *Bgl*II fragment from λCY46 cloned into the *Bgl*II site of pll99. The fragments used as hybridisation probes are red boxes.

of CYFl1-M28 (PN2303) and CYFl1-M61 (PN2301) were transformed with 5 μg of pCY40, pCY41 or pll99 and transformants selected on RG media containing 200 μg/mL of geneticin (Section 2.9). Ten arbitrarily selected geneticin-resistant colonies from each transformation were screened for intact copies of the *ltmM* or *paxM* genes by PCR amplification with primers lol20 and lol54, or paxMNotI and ggpps18, respectively. A subset of transformants that tested positive with the PCR screen were used in Southern analysis to determine the integration patterns and relative copy numbers of pCY40 and pCY41 (Fig. 3.27) using as probes either an *ltmM* or *paxM* fragment.

Southern hybridisation data showed that the four PN2303 transformants containing the *ltmM* gene on plasmid pCY40 have at least one - three copies of the *ltmM* fragment based on the comparison with the wild-type hybridisation signal (Fig. 3.27A). The PN2303 transformants containing the *paxM* gene on the pCY41 plasmid all contain more than one *paxM* copy with PN2303-paxM7 containing an extremely large *paxM* copy number (Fig. 3.27B, lane 7). The PN2301 transformants containing the *ltmM* gene on plasmid pCY40 range from approximately one copy to an extremely high copy number (Fig. 3.27A).

3.6.4 Lolitrem analysis of mutant associations containing *ltmM* or *paxM*

Perennial ryegrass seedlings were inoculated with the selected transformants (Section 2.10), as described above (Section 3.6.2). Inoculated plants were grown for approximately six weeks and then screened for endophyte infection using an endophyte-specific polyclonal antibody (Section 2.10). Two endophyte-infected plants for each selected transformant were maintained for lolitrem analysis. The rate of infection (Table 3.8), determined once the plants had established reasonable growth, shows that each transformant had similar rates of infection to those of wild-type Fl1. However, the infection rates for these inoculations were much higher than those previously examined (Section 3.6.2). The plants containing the transformants were screened for lolitrem production mid-spring (2003) and mid-summer (2004) (Fig. 3.28, Table 3.8). All plants that contained PN2303 transformed with pCY40 containing the *ltmM* gene were lolitrem positive. Of the plants that contained PN2303 transformed with pCY41 containing the *paxM* gene, only one, 2303-paxM7, was lolitrem positive complementing the *ltmM* defect. The extremely high *paxM* copy number in this strain is the most likely

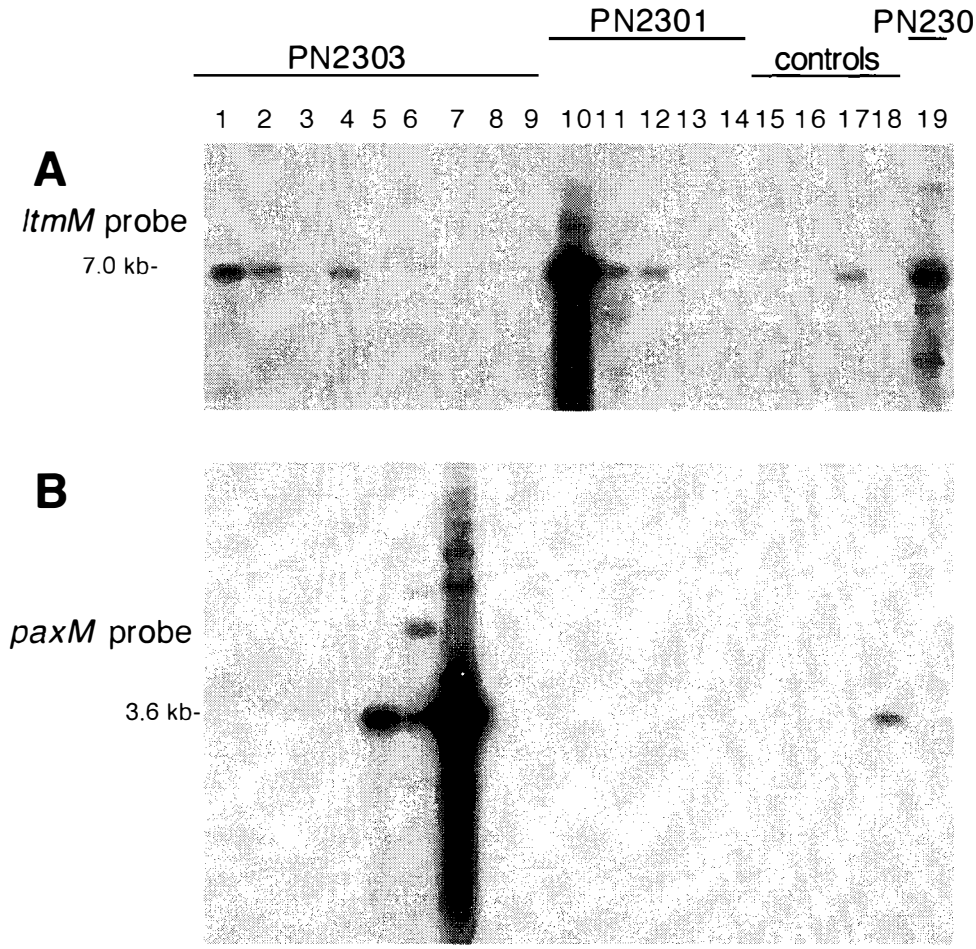


Figure 3.27 Southern analysis of the *ltmM* complementation transformants

Southern analysis to determine the relative copy number of the *ltmM* and *paxM* genes in the PN2303 and PN2301 *ltmM* deletion mutants. Blot (**A**) was hybridised with a ^{32}P-labelled *ltmM* fragment amplified with primers lol7 and lol28. Blot (**B**) was hybridised with a ^{32}P-labelled *paxM* fragment amplified with primers mono3 and mono4.

The 800 ng samples were digested and loaded as follows; Lane (1) PN2303-ltmM2; (2) PN2303-ltmM3; (3) PN2303-ltmM9; (4) PN2303-ltmM10; (5) PN2303-paxM2; (6) PN2303-paxM6; (7) PN2303-paxM7; (8) PN2303-pII99-6; (9) PN2303-pII99-8; (10) PN2301-ltmM6; (11) PN2301-ltmM7; (12) PN2301-ltmM10; (13) PN2301-pII99-4; (14) PN2301-pII99-10; (15) PN2303; (16) PN2301; (17) *E. festucae* FI1; (18) *P. paxilli*; (19) PN2301-ltmM3. Lanes 1-4, 8-17 and 19 were digested with *Xho*I. Lanes 5-7 and 18 were digested with *Bgl*II. Lanes 1-4 are PN2303 (CYFI1-M28) transformants containing pCY40 (the *ltmM* gene). Lanes 5-7 are PN2303 transformants containing pCY41 (the *paxM* gene). Lanes 10-12, and 19 are PN2301 (CYFI1-M61) transformants containing pCY40 (the *ltmM* gene). Lanes 8, 9, 13 and14 are PN2303 and PN2301 transformants, respectively, containing pII99.

Table 3.8 Rates of infection and lolitrem production for the associations containing the complementation strains

Strain	Plant association numbers	Integrating plasmid[1]	Infection Rate[2] (%)	plant/associations assayed	Lolitrem[3] (ppm)	Lolitrem[4] (ppm)
2303-ltmM2	G1167-68	pCY40	60	2	2.4-2.8	5.9*
2303-ltmM3	G1169-71	pCY40	59	3	2.7-3.0	NT
2303-ltmM9	G1172-73	pCY40	75	2	1.0-5.5	NT
2303-ltmM10	G1174-75	pCY40	71	2	0.4-0.8	1.2*
2303-paxM2	G1180-81	pCY41	63	2	0	0
2303-paxM4	G1182-83	pCY41	53	2	0	0
2303-paxM7	G1184-85	pCY41	33	2	1.4	8.8-12.9
2303-pII99-6	G1176-77	pII99	52	2	0	NT
2303-pII99-8	G1178-79	pII99	50	2	0	NT
2301-ltmM3	G1186-87	pCY40	62	2	0	NT
2301-ltmM6	G1188-89	pCY40	81	2	0	NT
2301-ltmM7	G1190-91	pCY40	81	2	0	NT
2301-ltmM10	G1192-93	pCY40	47	2	0	NT
2301-pII99-4	G1194-95	pII99	57	2	0	NT
2301-pII99-10	G1196-97	pII99	64	2	0	NT
CYFI1-M28	G1200-01	Δ*ltmM* (pCY39)	62	2	0	NT
CYFI1-M61	G1202-03	Δ*ltmMG* (pCY39)	67	2	0	NT
FI1	G1198-99	NA	82	2	3.7-4.0	10.7*
Endophyte free	G1204-05	NA	NA	2	0	NT

[1]The plasmid pCY40 contained the *ltmM* gene, pCY41 contained the *paxM* gene, pII99 was the control plasmid, pCY39 was the deletion construct Δ*ltmM* = *ltmM* insertional replacment, Δ*ltmMG* = extended deletion mutant, wt = wild-type.

[2]Infection rates were determined as a percentage of endophyte infected plants from the plants initially inoculated.

[3]The first lolitrem analysis was performed late October (2003). HPLC traces are shown in Fig. 3.28.

[4]Alkaloid analysis was performed mid-summer (January, 2004).

*Only one plant was tested.

NA = not applicable, NT = not tested.

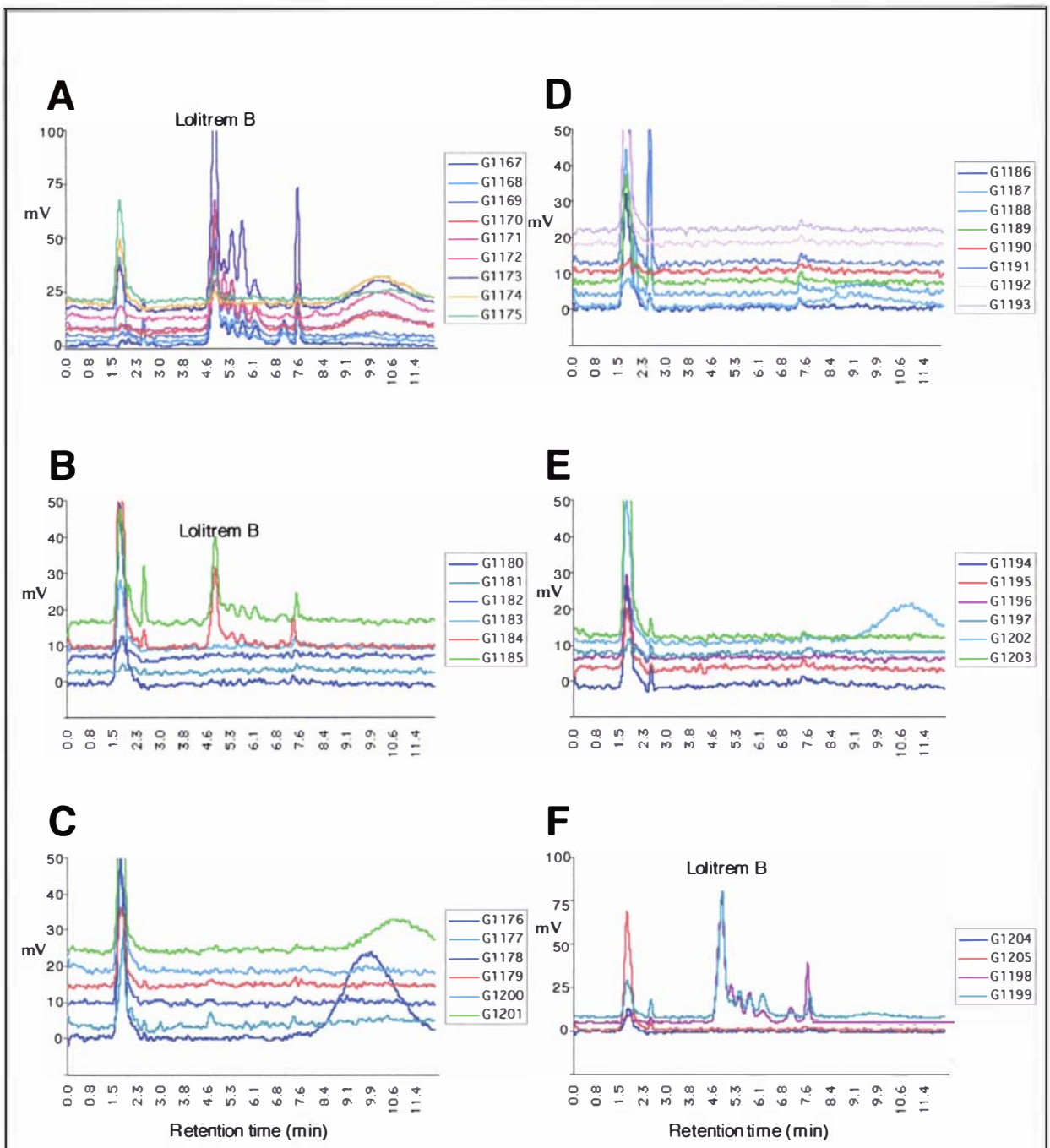


Figure 3.28 HPLC analysis for lolitrem B in the *ltmM*-deletion mutants complemented with *ltmM* or *paxM*

HPLC analysis of indole-diterpene extracts from pseudostems of endophyte-infected Nui perennial ryegrass. Extracts of Nui perennial ryegrass infected with (**A**) PN2303 transformed with pCY40 (containing the *ltmM* gene), (**B**) PN2303 transformed with pCY41 (containing the *paxM* gene), (**C**) PN2303 containing pII99 (G1176-G1179) or PN2303 (G1200-G1201), (**D**) PN2301 transformed with pCY40, (**E**) PN2301 containing pII99 (G1194-G1197) or PN2301 (G1202-G1203), (**F**) FI1 (G1198-G1199) or uninfected (G1204-1205).

PN2303 was the CYFI1-M28 mutant containing a deleted *ltmM* gene. PN2301 was the CYFI1-M61 mutant containing a deletion of *ltmM*, *ltmG* and DNA upstream of *ltmG*. The peaks eluting later than lolitrem B contain lolitrems A and E.

cause for lolitrem production rather than the presence of common regulatory elements within the *paxM* and *ltmM* promoters. As expected, none of the plants that contained the PN2301 transformants with pCY40 was able to produce lolitrems as both *ltmM* and *ltmG* are deleted in this strain.

3.6.5 Complementation of *P. paxilli paxM* deletion

Complementation was used to test whether *ltmM* would complement the *P. paxilli paxM*-deletion mutant LMM100 when fused to a *paxM* promoter. The *paxM* promoter was selected to control the *ltmM* gene so that the regulation would be similar to the other *pax* genes thereby optimising paxilline production, and also because of the presence of a convenient *Nco*I site at the translation start of the *paxM* gene. An 894-bp *paxM* promoter fragment, containing ~850 bp of promoter from the *paxM* ATG, was amplified with Expand High Fidelity PCR system using primers CY9 and CY11, digested with *Bam*HI and cloned into the *Bam*HI site of pUC19, resulting in plasmid pPN1851 (Fig. 3.29). The *ltmM* genes from Lp19 and Fl1 were amplified with Platinum *Pfx* (Invitrogen) using primers lol284 and lol285 containing *Nco*I and *Sst*I sites respectively. The *ltmM* PCR product comprised of 1833 bp of coding sequence and 170 bp of 3' untranslated region. The *ltmM* genes were fused to the *paxM* promoter at the ATG translational start site using the restriction enzymes *Nco*I and *Sst*I (Figure 3.29). The translational fusion that results in creating an *Nco*I site in the *ltmM* gene causes a single base change where the second codon of *ltmM* has a conservative replacement of threonine in the wild-type gene to alanine in the fused gene. The sequence of the resulting clones, pCY54 and pCY55 containing Lp19 and Fl1 *ltmM* respectively, contained no PCR amplification errors. Protoplasts of LMM100 were transformed with pII99, pCY40, pCY41, or co-transformed with pII99 and either pCY54 or pCY55, and transformants were selected on geneticin. Approximately 5-10 stable *P. paxilli* LMM100 transformants were colony-purified and subsequently screened by TLC analysis for their ability to produce paxilline (Fig. 3.30).

TLC analysis of the wild-type *P. paxilli* indole-diterpenoid extract shows an intense green band that has the same Rf as paxilline (Fig. 3.30). The wild-type sample also contains other green and yellow bands that could be indicative of other indole-diterpene compounds, such as paspaline and 13-desoxypaxilline (Fig. 3.30). The LMM100 *paxM*-deletion mutant, used as the recipient in the transformations, was unable to

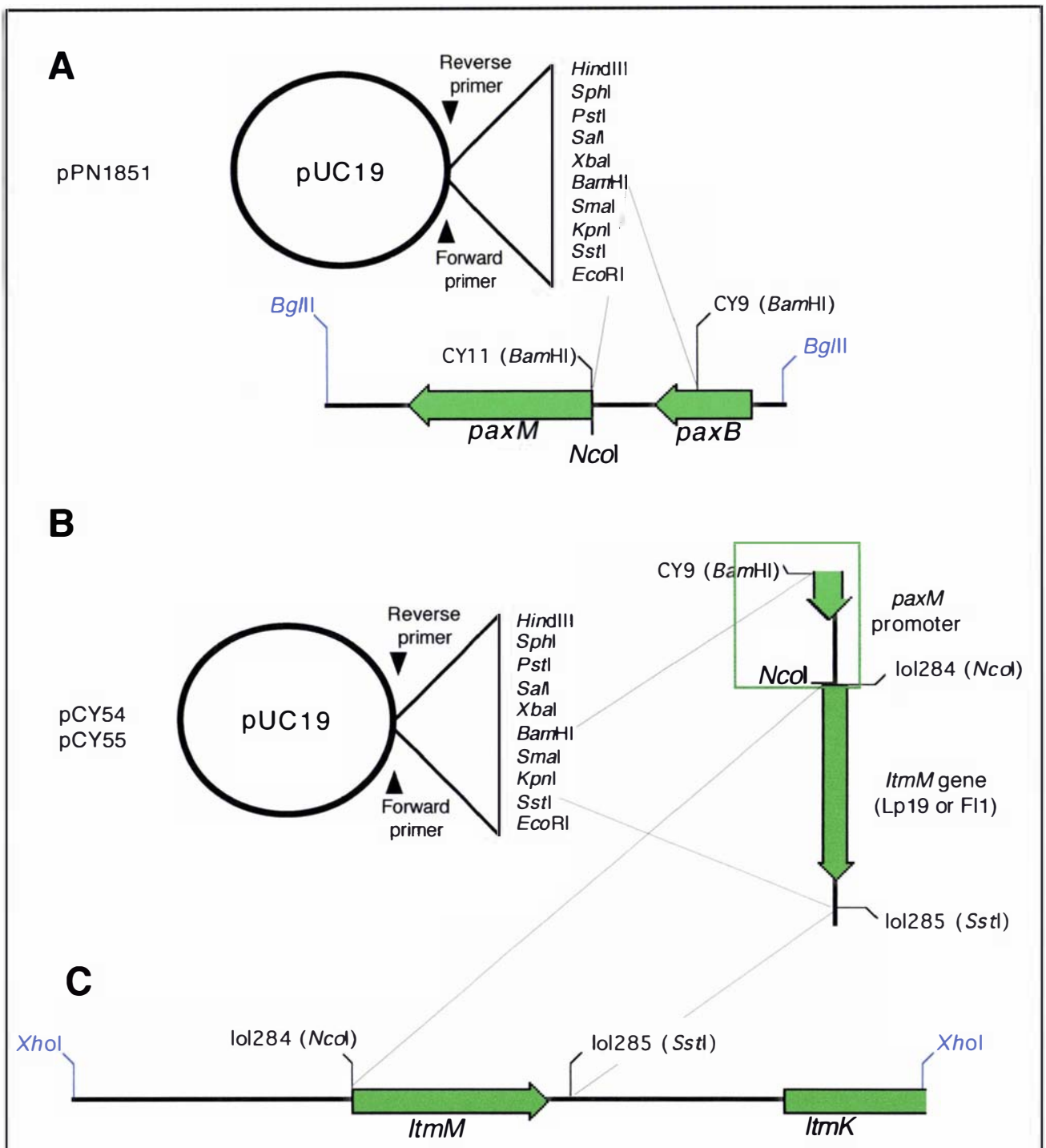


Figure 3.29 Constructs for complementation of the *paxM*-deletion mutant

The constructs for complementation of the *paxM*-deletion mutant. (**A**) The pPN1851 construct. The *paxM* promoter was amplified with primers CY9 and CY11, digested with *Bam*HI and cloned into the *Bam*HI site of pUC19. (**B**) The pCY54 and pCY55 constructs. The *ltmM* genes from Lp19 (pCY54) and Fl1 (pCY55) were amplified with primers lol284 and lol285, digested with *Nco*I and *Sst*I and subsequently cloned into pPN1851. The *paxM* promoter is surrounded by a green box. (C) The 7 kb XhoI fragment used in pCY40 (Fig. 3.26A).

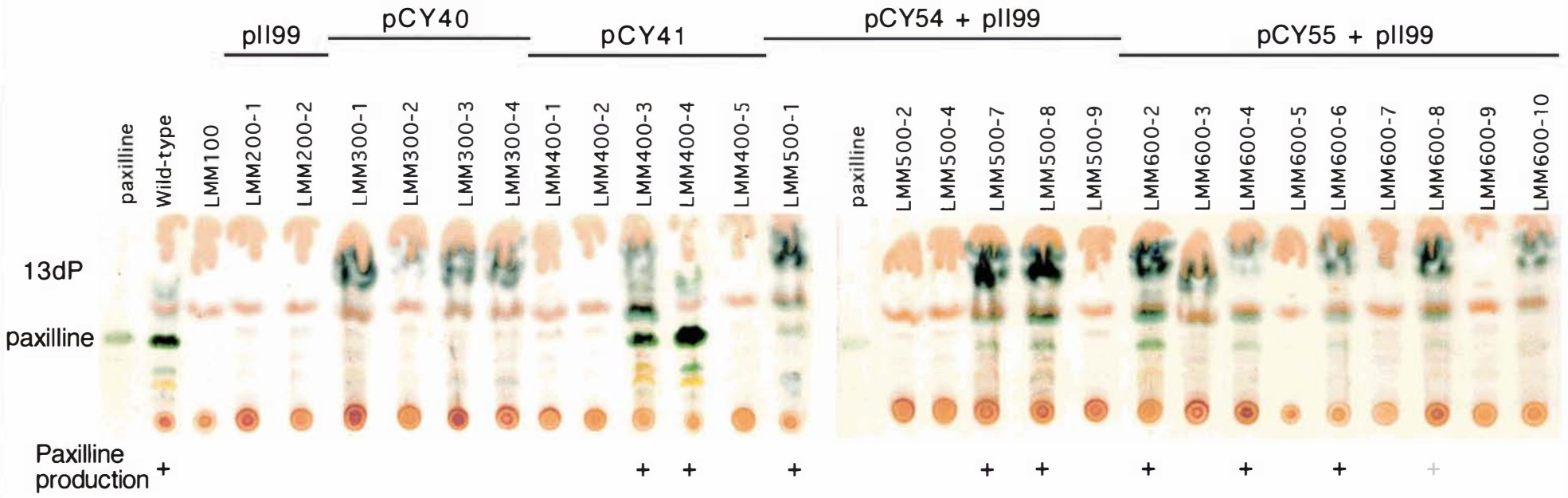


Figure 3.30 TLC analysis of the *paxM* complementations

TLC analysis of the *paxM* mutant, LMM100, containing complementation constructs pCY40, pCY41, pCY54 and pCY55. Indole-diterpenes were extracted from mycelia (Section 2.11.5) grown for 7 days in CDYE + TE (Section 2.2.2, 2.2.7) at 28^{o}C. The transformants analysed included LMM100 containing pII99 with the geneticin selectable marker; pCY40 with *ltmM* under the control of its native promoter; pCY41 with *paxM* under the control of its native promoter; pCY54 with the Lp19 *ltmM* gene under the control of the *paxM* promoter; pCY55 with the Fl1 *ltmM* gene under the control of the *paxM* promoter. The + indicates a green band of the same Rf as paxilline, while + indicates possible paxilline production that was confirmed by HPLC analysis (Appendix 5.4). 13dP = the approximate Rf of 13 desoxypaxilline and paspaline.

produce any indole-diterpenes (Fig. 3.30). The LMM100 transformants, LMM200-1 and LMM200-2, that contained the pII99 vector, were unable to complement the *paxM* deletion and are therefore paxilline negative (Fig. 3.30). LMM100 transformants that contained the plasmid pCY40 with the Lp19 *ltmM* gene under the control of the native Lp19 *ltmM* promoter were unable to complement the *paxM* deletion (Fig. 3.30; samples LMM300-1 to -4). Two of the five LMM100 transformants that contained the endogenous *paxM* gene on pCY40 were able to complement the *paxM* deletion phenotype (Fig. 3.30; samples LMM400-1 to -5). Further analysis would be required to determine why three of the five transformants were unable to produce paxilline or other indole-diterpenes but could be due to the integration of the plasmid in a region of the genome that did not support expression of the transgene, such as within heterochromatin. TLC analysis showed that ~50% of the LMM100 transformants that contained the *ltmM* gene under the control of the *paxM* promoter had detectable levels of paxilline, confirming the *paxM* deletion was complemented by *ltmM* (Fig. 3.30). As expected, no differences were observed between the Lp19 and FII *ltmM* genes in the complementation. The single amino acid change of the second *ltmM* codon in pCY54 and pCY55 had no effect on *paxM* complementation. The TLC results of selected samples were confirmed by HPLC analysis (Appendix 5.4). However, additional analysis will be required to determine whether the other green bands that are present in many of the samples (Fig. 3.30) are indole diterpenes.

The LMM100 transformants were arbitrarily selected as geneticin-resistant colonies and were not screened for intact copies of the integrating plasmids. Therefore, the frequency of successful co-transformation was determined as a percentage of geneticin resistant transformants that also had an ability to produce indole-diterpenes. The percentage of LMM100 transformants co-transformed with pII99 and either pCY54 or pCY55 was ~50%.

3.7 The regions flanking the three *ltm* genes

The hybridisation of the CY28 fragment to the λGEM12 Lp19 genomic library (Section 3.3) resulted in the isolation of one major clone, λCY218, of which the genomic position is skewed to the right of *ltmG* (Fig. 3.11, 3.31). Five clones were in fact isolated but λCY218 contained the largest amount of unique sequence. The remaining

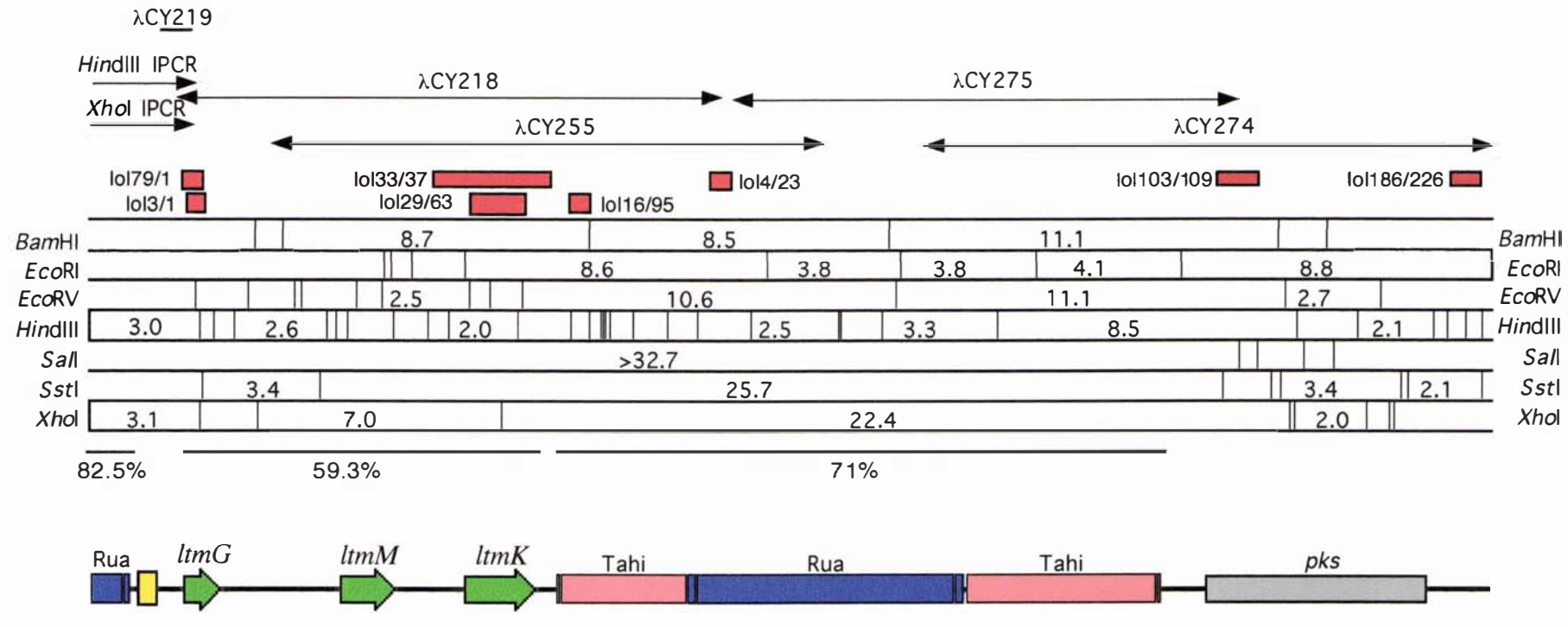


Figure 3.31 A physical map of the Lp19 *ltm* cluster

A physical map of the *ltm* cluster. The fragment sizes in the restriction enzyme map are stated in kb. The lambda clones used for generating the sequence are shown as double-ended arrows above the region they span. The genes are shown as green arrows, the *pks* pseudogene is a grey box, the yellow box is a microsatellite sequence, the retrotransposon Tahi is a light red box, the retrotransposon Rua is a blue box. Fragments used for Southern hybridisation analysis are shown as red boxes above the restriction enzyme map with the amplification primers adjacent to the box. The AT content is shown as a percentage with a line to indicate the region analysed.

four clones contained smaller lengths of λCY218 or they contained rearrangements and did not conform to expected hybridisation data or sequence analysis. This skewed and underrepresented pattern was in contrast to the *ggs1* locus where approximately 22 kb was represented by at least nine lambda clones (Section 3.2.1). Sequence analysis of λCY218 revealed a total of three genes, *ltmG*, *ltmM* and *ltmK*, flanked by AT-rich DNA. This AT-rich sequence contained very weak similarity to the reverse transcriptase region of retrotransposon sequences. The only detectable homology was to a ZAM retrovirus (accession number CAA04050; Leblanc, et al., 1997) from *Drosophila melanogaster* (E-value 0.84). Although the homology to the ZAM retrovirus was weak, signature sequences to reverse transcriptase and RNase H domains were identified within the translated sequence of the AT-rich sequence. However, the sequence lacked intact open reading frames, which suggested that *ltmG*, *ltmM* and *ltmK* could be flanked by degenerate retrotransposons.

As the lolitrem B biochemical pathway is predicted to contain a gene cluster of at least 8-12 genes (Section 1.10), attempts were made to isolate additional flanking sequence. Two approaches were used to isolate additional sequence, IPCR and library screening with probes used from the extreme ends of λCY218 and to *ltmK*.

3.7.1 Extending the left-hand flanking sequence

The hybridisation of a *ltmG* fragment, amplified with primers lol1 and lol3, to the Lp19 λGEM12 genomic library resulted in the isolation of one new clone, λCY219, but this clone was rearranged and contained only 1 kb of additional *ltmG* flanking sequence (Fig. 3.31).

Southern analysis of Lp19- and Fl1-digested genomic DNA hybridised with a *ltmG* fragment revealed a number of large hybridising fragments (Fig. 3.32A). However, two restriction enzymes, *Hin*dIII and *Xho*I, generated smaller hybridising fragments of a size suitable for isolation of additional flanking sequence from the left-hand side of *ltmG* by IPCR. IPCR with the self-ligated *Hin*dIII and *Xho*I digests of Lp19 genomic DNA (Fig. 3.32B) resulted in the isolation of an additional ~ 1.8 kb of sequence (Fig. 3.31, 3.32C). Sequence analysis of this region in Lp19 revealed a microsatellite consisting of a core TAATG sequence repeated ~104 times. To the left of the microsatellite sequence the sequence was very AT-rich (82.5% AT) with sequence

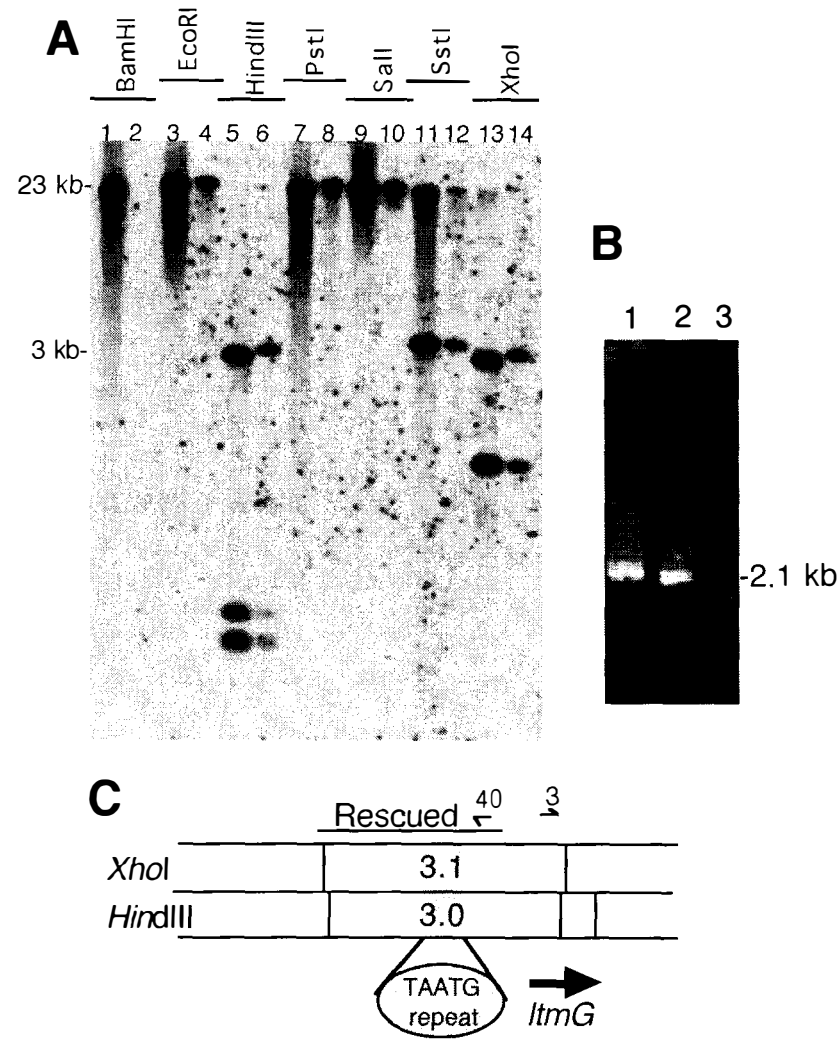


Figure 3.32 IPCR to rescue left hand flanking sequence

(**A**) Southern analysis to determine restriction enzymes suitable for IPCR. Lane (1, 3, 5, 7, 9, 11 and 13) 2 μg of Lp19 genomic DNA; (2, 4, 6, 8, 10, 12 and 14) 2 μg of Fl1 genomic DNA, digested as indicated. The Southern blot is hybridised with a ^{32}P-labelled *ltmG* probe amplified with primers lol79 and lol27. The Fl1 *Bam*HI digest floated out of the well during loading.

(**B**) IPCR with (1) *Xho*I, (2) *Hin*dIII digested then self-ligated Lp19 DNA. PCR amplification with primer pair lol40 and lol3.

(**C**) A genomic map showing the region rescued by IPCR, the location of the primers and the orientation *ltmG*.

The following PCR reaction conditions were used: 50 ng self-ligated digested genomic DNA, 1x Expand HiFi buffer (Roche), 50 μM each dNTP, 200 nM each primer, 0.4 μL Expand HiFi (Roche) in a reaction volume of 50 μL. The PCR amplification conditions were as follows: 94^{o}C for 2 min, followed by 10 cycles of 94^{o}C for 30 s, 55^{o}C for 30 s, 68^{o}C for 3 min 30 s, then 20 cycles of 94^{o}C for 30 s, 55^{o}C for 30 s, 68^{o}C for 3 min 30 s with a 10 s increment/extension cycle, finally one cycle of 68^{o}C for 7 min.

similarities to retrotransposons. However, once again there was no evidence of intact open reading frames.

3.7.2 Extending the right hand flanking sequence

The hybridisation of a fragment amplified with lol4 and lol23 to the Lp19 λGEM12 library resulted in over 100 hybridising clones per library plate containing ~8000 pfu. Twelve arbitrarily selected clones were further analysed. However, only seven, λCY224, λCY225, λCY226, λCY227, λCY229, λCY230, and λCY231 remained positive after three rounds of library screening. The clones were sequenced with primers, SP6 and T7, which anneal to the lambda arms. Of the seven positive clones two sequences, λCY224 and λCY230 were identical, one clone, λCY225, gave poor sequence data when used as a template and λCY231 was rearranged and contained lambda sequence. The sequence of these clones was extended using sequence-specific primers. However, extending the sequence was difficult because of the AT content and repetitive nature of the clones. None of these clones directly overlapped with λCY218 but they did contain regions within them that were near identical. This will be discussed further in Section 3.8. BLASTX analysis of λCY224, λCY227 and λCY229 had database matches to retrotransposon sequences while λCY226 had no database match.

The library screen with an *ltmK* fragment, amplified with primers lol33 and lol37, resulted in the hybridisation of 28 positive clones. Only one of these clones, λCY255, extended beyond the λCY218 sequence (Fig. 3.31). The sequence of the λCY255 clone was extended by direct sequencing adding an additional 2.9-kb to the *ltm* cluster. However, the AT content remained high with no obvious open reading frames. A match to the retrotransposon Skippy from *Fusarium oxysporum* (E-value 6e-19) sequence suggested the likelihood of a retrotransposon platform downstream from *ltmK*.

3.7.3 Using Fl1 to jump across the retrotransposon platform

As discussed earlier (Section 3.6) *N. lolii* strain Lp19 and *E. festucae* strain Fl1 are almost identical throughout the entire *ltm* region. However, Southern hybridisation with an *ltmK* fragment to genomic digests of Lp19 and Fl1 revealed a different banding pattern surrounding the *ltmK* fragment (Fig. 3.33). The unique banding pattern of Fl1 suggested that the retrotransposon sequence present in Lp19 was absent in Fl1.

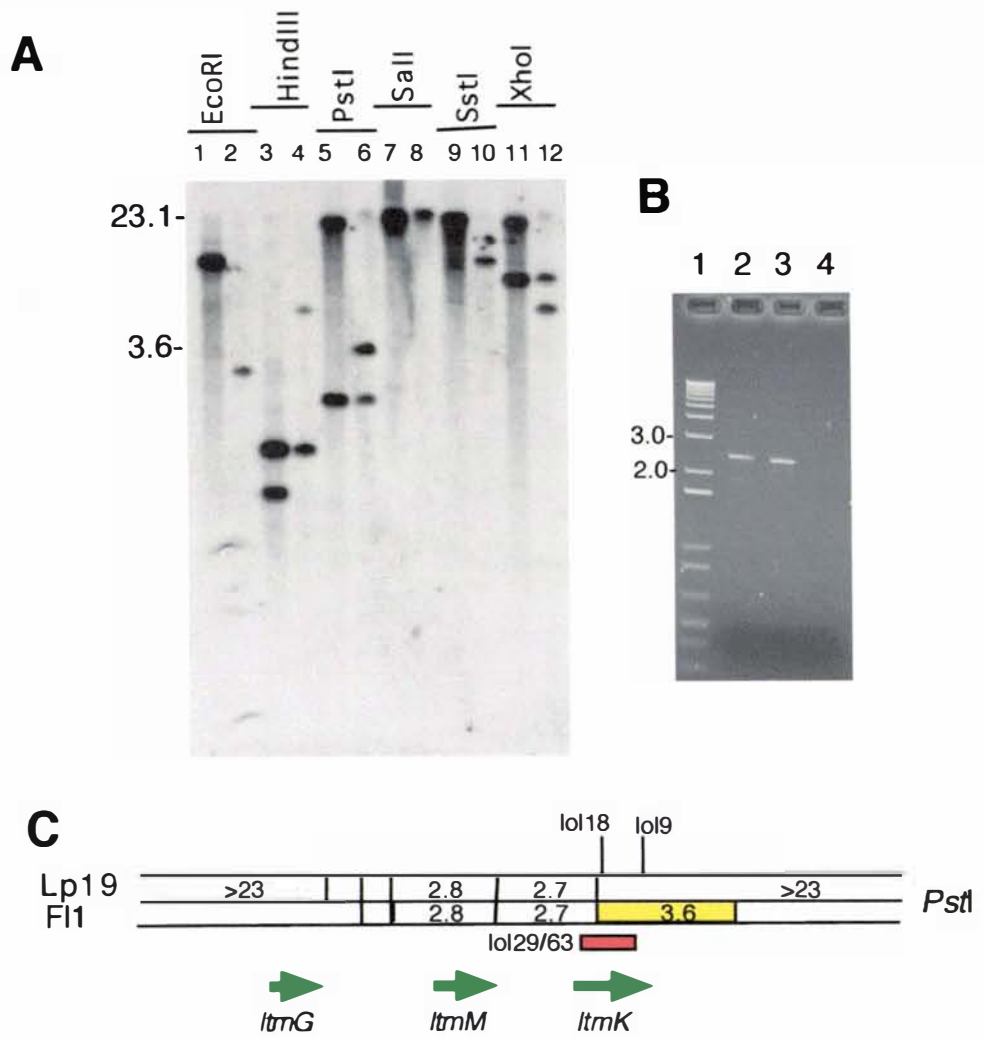


Figure 3.33 Restriction enzyme map differences between Lp19 and Fl1 at the *ltm* junctions

(**A**) Southern analysis to identify restriction enzymes suitable for IPCR. Autoradiograph of a hybridisation with a ^{32}P-labelled *ltmK* fragment (shown as a red box in C) amplified with primers lol29 and lol63. Lane (1, 3, 5, 7, 9 and 11) 2 μg of Lp19 digested genomic DNA, (2, 4, 6, 8, 10 and 12) 2 μg of Fl1 digested genomic DNA.

(**B**) IPCR with (2) Fl1, (3) E189 *Pst*I-digested then self-ligated DNA amplified with primers lol18 and lol9. Lanes (1) 1 kb + ladder (Invitrogen) and (4) water control.

(**C**) Physical maps of the restriction enzyme *Pst*I from Lp19 and Fl1 showing the region rescued by IPCR (yellow box) and the location of the primers lol18 and lol9 used in the IPCR reaction.

The following PCR reaction conditions were used: 50 ng digested self-ligated genomic DNA, 1x *Taq* polymerase buffer (Roche), 100 μM each dNTP, 200 nM each primer, 2 U *Taq* polymerase (Roche) in a reaction volume of 50 μL. The PCR amplification conditions were as follows: 94^{o}C for 2 min, followed by 30 cycles of 94^{o}C for 30 s, 58^{o}C for 30 s, 72^{o}C for 2 min 30 s, then one cycle of 72^{o}C for 5 min.

Analysis of the hybridisation data indicated the 3.6 kb *Pst*I hybridising fragment present in Fl1 would contain ~2.3 kb of unique sequence when isolated by IPCR with *Pst*I-digested then self-ligated Fl1 genomic DNA. IPCR with primers lol18 and lol9 resulted in the amplification of a 2.5-kb band (Fig. 3.33B). Due to the presence of non-specific bands in the PCR amplification, the product was cloned into pGEM-T easy then sequenced. Sequence analysis revealed the presence of a putative gene, with strong BLASTX matches to polyketide synthases. The best match was to AN9217 from *Aspergillus nidulans* with an E-value of 9e-10. This sequence was extended a further 1.5-kb with IPCR using as a template Fl1 *Bam*HI-digested then self-ligated genomic DNA with primers lol103 and lol13 (Fig. 3.34). BLASTX analysis with the additional sequence resulted in a more significant match to *A. nidulans* AN9217 with an E-value of 3e-83. There is no predicted role for a polyketide synthase in indole-diterpene biosynthesis (Section 1.10).

3.7.4 The Lp19 *pks* is separated from *ltmK* by a retrotransposon platform

To determine whether the *pks* gene was contained on the same *N. lolii* Lp19 chromosome as the *ltm* cluster a chromosome separation of Lp19 was hybridised with a *pks* fragment (Fig. 3.35). The *pks* fragment hybridised to the same 2.8 Mb Lp19 chromosome as the *ltmG* fragment (Fig. 3.35).

The mapped *E. festucae* Fl1 sequence (Fig. 3.34) revealed that the only *Sal*I sites across the *ltm* cluster were contained within the region rescued by IPCR. Restriction enzyme analysis of the *N. lolii* Lp19 and *E. festucae* Fl1 *ltm* region (Fig. 3.31; 3.34) showed that the *ltmK* gene was contained on *Sal*I and *Sst*I fragments greater than 23 kb. To determine whether *ltmK* and *pks* hybridised to the same large fragments, pulse-field gel electrophoresis of *Pst*I, *Sal*I and *Sst*I digests was performed. Both the *ltmK* and *pks* fragments hybridised to the same ~73 kb *Sal*I and 26 kb *Sst*I fragments (Fig. 3.36) in the *Neotyphodium* strains Lp1 and Lp5.

Using the *pks* fragment amplified with primers lol103 and lol109 as a probe, the Lp19 *pks* sequence was isolated from the Lp19 λGEM-12 genomic library. Eleven positive clones covering approximately 21 kb of genomic DNA sequence were isolated. λCY275 and λCY274 were the most diverse overlapping clones, with λCY275 overlapping with λCY255 by ~2.9 kb (Fig. 3.31).

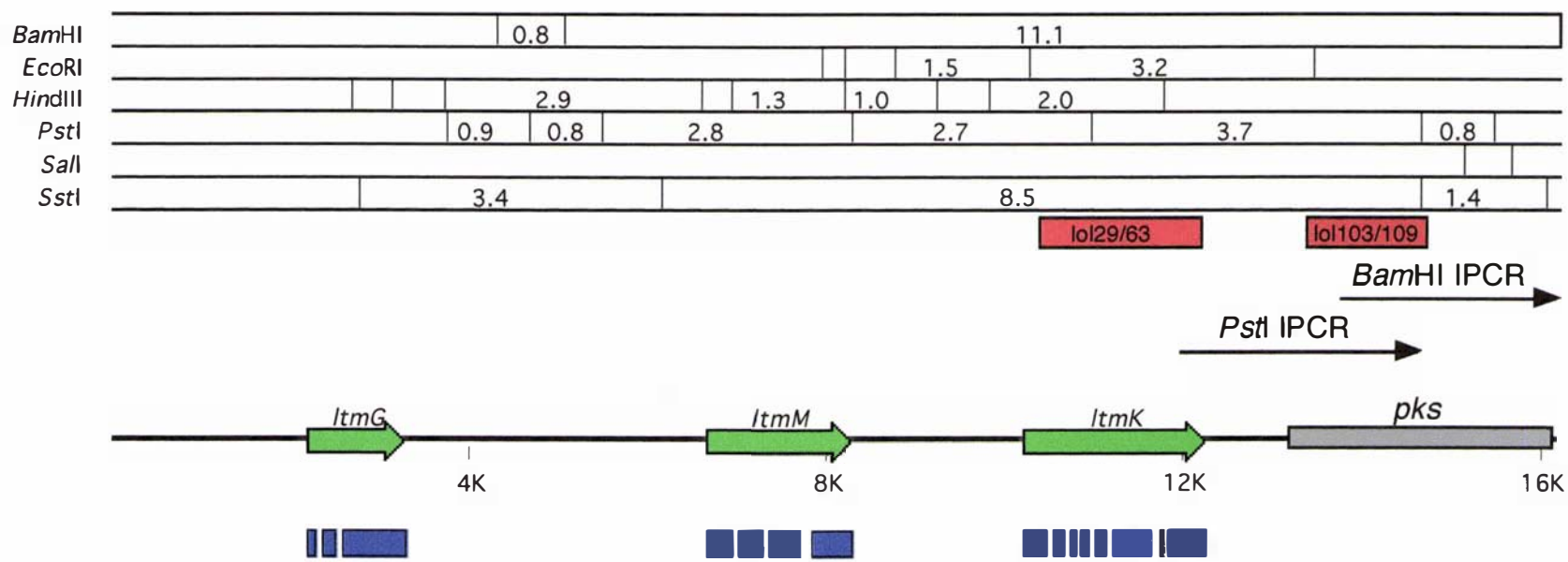


Figure 3.34 Physical and genetic map of the *E. festucae*, Fl1, *ltm* locus

Physical and genetic map of the Fl1 *ltm* locus showing the region used to isolate sequence flanking *ltmK* by IPCR. The *ltm* genes are green arrows, the *pks* pseudogene is a grey box. The exons for the *ltm* genes are blue boxes under the gene. The regions rescued by IPCR are black arrows. The remaining Fl1 sequence was generated by PCR amplification. The PCR products were purified and sequenced directly with sequence-specific primers. The fragments used for hybridisation are shown as red boxes.

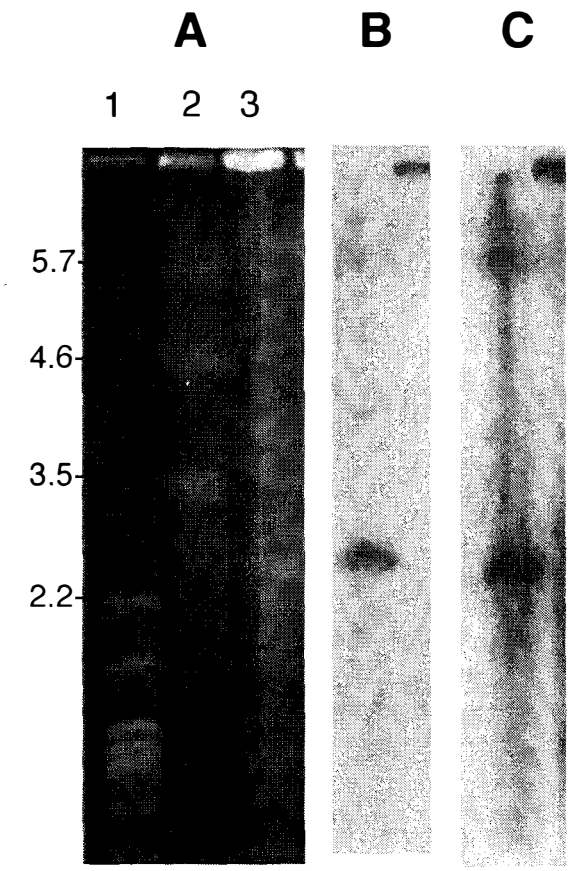


Figure 3.35 Southern analysis of the *pks* pseudogene

Separation of the Lp19 chromosomal DNA by pulse-field electrophoresis. (**A**) Ethidium bromide-stained gel of the Lp19 chromosomal DNA separations. Lane (1) *S. cerevisiae* chromosomal DNA standard (BioRad), (2) *S. pombe* chromosomal DNA standard (BioRad), (3) Lp19 chromosomal DNA. The DNA standards are shown as Mb. The electrophoresis conditions were 0.7% (w/v) chromosomal grade agarose (Biorad) in 0.5x TBE at 60 volts with the following switching times; 120 s for 24 h, 450 s for 20 h, 1500 s for 72 h, 2100 s for 60.5 h at 14^{o}C.

(**B**) Autoradiograph of Lp19 chromosomal DNA from (**A**) hybridised with a ^{32}P-labelled *ltmG* fragment amplified with primers lol3 and lol1.

(**C**) Autoradiograph of Lp19 chromosomal DNA from (**A**) hybridised with a ^{32}P-labelled *pks* fragment amplified with primers lol103 and lol109.

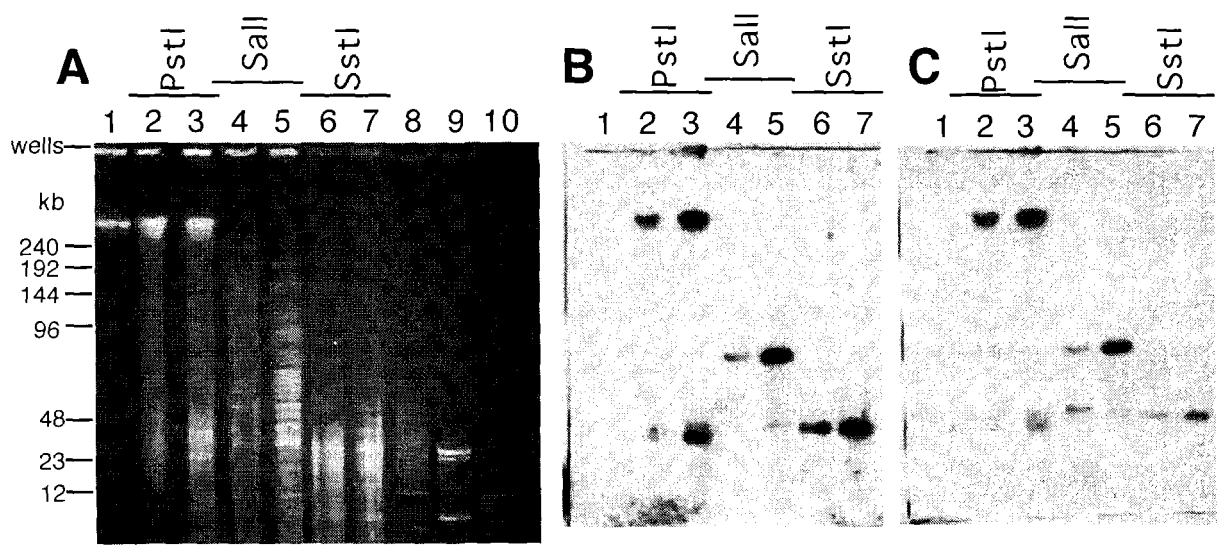


Figure 3.36 Linkage of *ltmK* and *pks*

Southern analysis to show linkage of *ltmK* and *pks* sequences. Lane (1) Lambda ladder (Biorad); *Pst*I-digested DNA from (2) Lp1, (3) Lp5; *Sal*I-digested DNA from (4) Lp1, (5) Lp5; *Sst*I-digested DNA from (6) Lp1, (7) Lp5; (8) 8-48 kb ladder (Invitrogen), (9) Lambda DNA *Hin*dIII-digested, (10) 1 kb plus ladder (Invitrogen).

(**A**) Ethidium bromide stained gel of pulse-field electrophoresis separations of *Pst*I-, *Sal*I- and *Sst*I-digested Lp1 and Lp5 genomic DNA. The electrophoresis conditions were 1% (w/v) Molecular biology grade agarose (Roche) in 0.5 x TBE at 200 volts with ramped switch times of 2 - 10 s for 18.4 h and 10-16 s for 4.6 h at 14^{O}C.
(**B**) Autoradiograph of (A) hybridised with a ^{32}P-labelled *ltmK* fragment amplified with primers lol29 and lol63.
(**C**) Autoradiograph of (A) hybridised with a ^{32}P-labelled *pks* fragment amplified with primers lol103 and lol109.

The *pks* sequence was analysed by BLASTX and showed significant matches to polyketide synthases with the best database match to the characterised PKS1 from *Cochliobolus heterostrophus* (E-value e-151; Yang, et al., 1996a). The *N. lolii* Lp19 *pks* gene has been considered a pseudogene as it contains a number of frame shifts with no obvious intron splice sites. There is no detectable in-frame start methionine codon of the *pks* gene and the sequence is flanked by AT-rich DNA that is devoid of open reading frames. Transcripts from this pseudogene could not be detected.

3.8 Definition of the retrotransposon platform

Sequence analysis of the Lp19 region between *ltmK* and *pks* indicated that this region contained evidence of a retrotransposon platform (Fig. 3.31). This was supported by the presence of two copies of two independent long terminal repeats (LTR) thereby allowing the identification of two retrotransposons, Tahi and Rua (Maori for one and two).

The LTR sequences for Tahi are ~104 bp, and are flanked by 5-bp target site duplications of CGCGC. The Tahi element has been interrupted by the insertion of the Rua element. As mentioned previously, the RT and RNase H domains of the *pol* gene were detected by BLASTX matches to the available databases. Based on this analysis an uninterrupted Tahi element is predicted to be approximately 9.3 kb. The best BLASTX match of the complete Tahi element is to a reverse transcriptase region from *Oryza sativa* (accession number CAE03017; E-value 0.86). The significance associated with the BLASTX analysis is poor, which indicated that the sequence that remained was a relic of a retrotransposon.

The LTR sequences of Rua are ~272 bp and are flanked by 5-bp target site duplications of TCCTT. The approximate size of the complete Rua element is 7.9 kb. The complete Rua retrotransposon has a closest BLASTX match to the reverse transcriptase and integrase region of a retrotransposon from *Cryptococcus neoformans* (accession number EAL20356; E-value 8e-66). When the sequence of the Rua LTR was compared to the *N. lolii* Lp19 *ltm* cluster consensus, three Rua LTRs were identified. The third copy of a Rua LTR was found to be adjacent to *ltmG* (Fig. 3.31). However, this retro element is truncated with a putative *pol* gene immediately adjacent to the LTR. Although the

Fl1 strain lacks the combined Tahi and Rua platform found between *ltmK* and *pks* in Lp19, the Rua retrotransposon adjacent to *ltmG* in Lp19 is present in *E. festucae* Fl1 and is identical to the available Lp19 DNA sequence.

Based on the order and sequence of the motifs identified within the *pol* domain, both Tahi and Rua belong to the Class I Gypsy sub-family of retrotransposons (Kempken, Kück, 1998). However, the two retro elements have no significant nucleotide homology to one another and have different LTR sequences, suggesting they are independent elements.

When fragments from the Tahi and Rua elements and the AT-rich sequence flanking the *pks* were used as hybridisation probes against Southern blots containing *N. lolii* Lp19, *E. festucae* Fl1 and *E. typhina* E8 genomic digests, unique hybridisation patterns with each probe were obtained (Fig. 3.37). In both Lp19 and Fl1 strains, Rua is present in the highest copy number followed by the AT-rich probe and then Tahi. None of the probes hybridised to E8 genomic digests. Each probe results in different banding patterns between the Lp19 and Fl1 strains with the three restriction enzyme digests used. However, for each probe Lp19 and Fl1 contained conserved hybridising *Hin*dIII fragments that are present in multiple copies.

The LTR sequences of Tahi and Rua were used to search a database of sequences generated from this thesis. A Tahi element was identified in λCY226 (Section 3.7.2) and three additional Rua elements were identified, one each in λCY226 and λCY231 (Section 3.7.2) and one in *ltm* cluster 2 which will be presented in Section 3.9.1. An alignment of the LTR sequences using ClustalW showed that the nucleotide differences within the alignment were predominantly G and A or C and T (Fig. 3.38). This preferential conversion of G:C to A:T is a feature of *Neurospora crassa* known as repeat induced point mutation (RIP) (Section 1.11.2; Cambareri, et al., 1989). Of the 20-nucleotide differences seen in Tahi LTRs, 13 are G:C to A:T nucleotide transitions. The Rua LTR sequences have 82 G:C to A:T nucleotide transitions and 13 others. To ascertain whether the transitions observed were similar to that of RIP, the preference of nucleotide residue that flanked the G:C to A:T transitions was determined. However, transitions that occurred in tandem were excluded from the analysis because the order of transitions was unable to be determined. The Tahi LTR sequences had 75% transitions that were unequivocally considered to have a CpA consensus, as there were

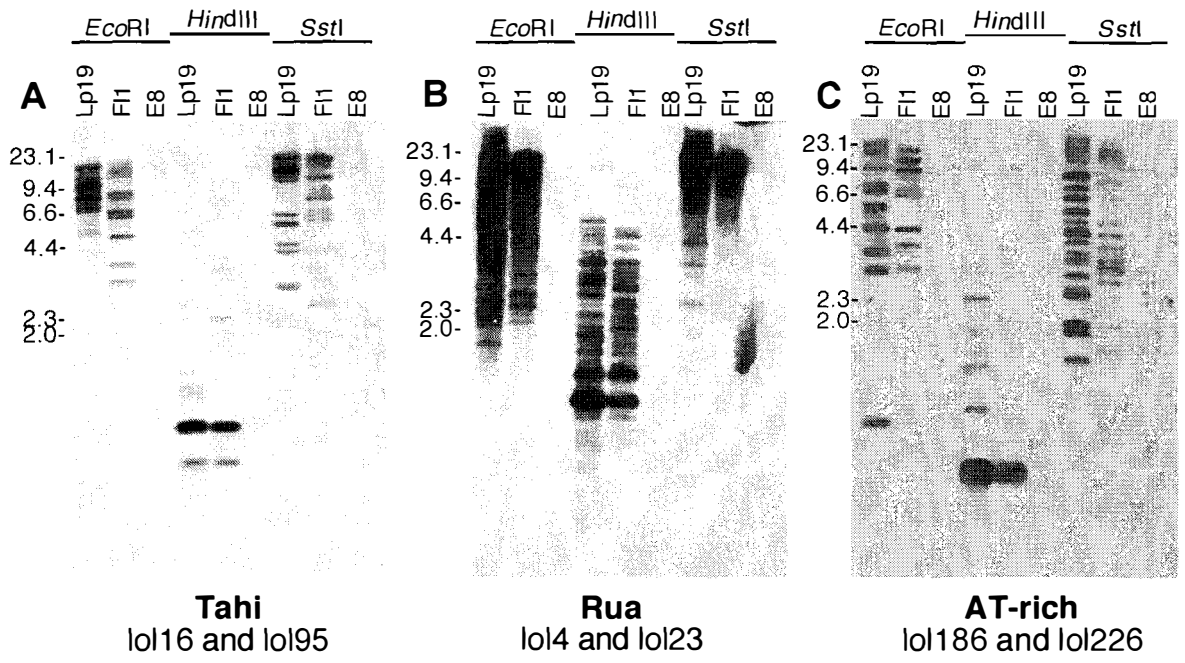


Figure 3.37 Southern analysis to determine copy number of Tahi, Rua and the AT-rich sequence

Southern analysis to determine the copy number of the retro-elements. Tahi and Rua, and the AT-rich sequence flanking *ltm* cluster 1. *N. lolii* Lp19, *E. festucae* Fl and *E. typhina* E8 genomic DNA (2 μg) was digested with the restriction enzymes, *Eco*RI, *Hin*dIII and *Sst*I. Size standards are in kb. Probe positions within the ltm cluster are shown in Fig. 3.31

(**A**) Autoradiograph of hybridisation with a ^{32}P-labelled Tahi fragment amplified with primers lol16 and lol95.
(**B**) Autoradiograph of hybridisation with a ^{32}P-labelled Rua fragment amplified with primers lol4 and lol23.
(**C**) Autoradiograph of hybridisation with a ^{32}P-labelled fragment of the AT-rich region amplified with primers lol186 and lol226.

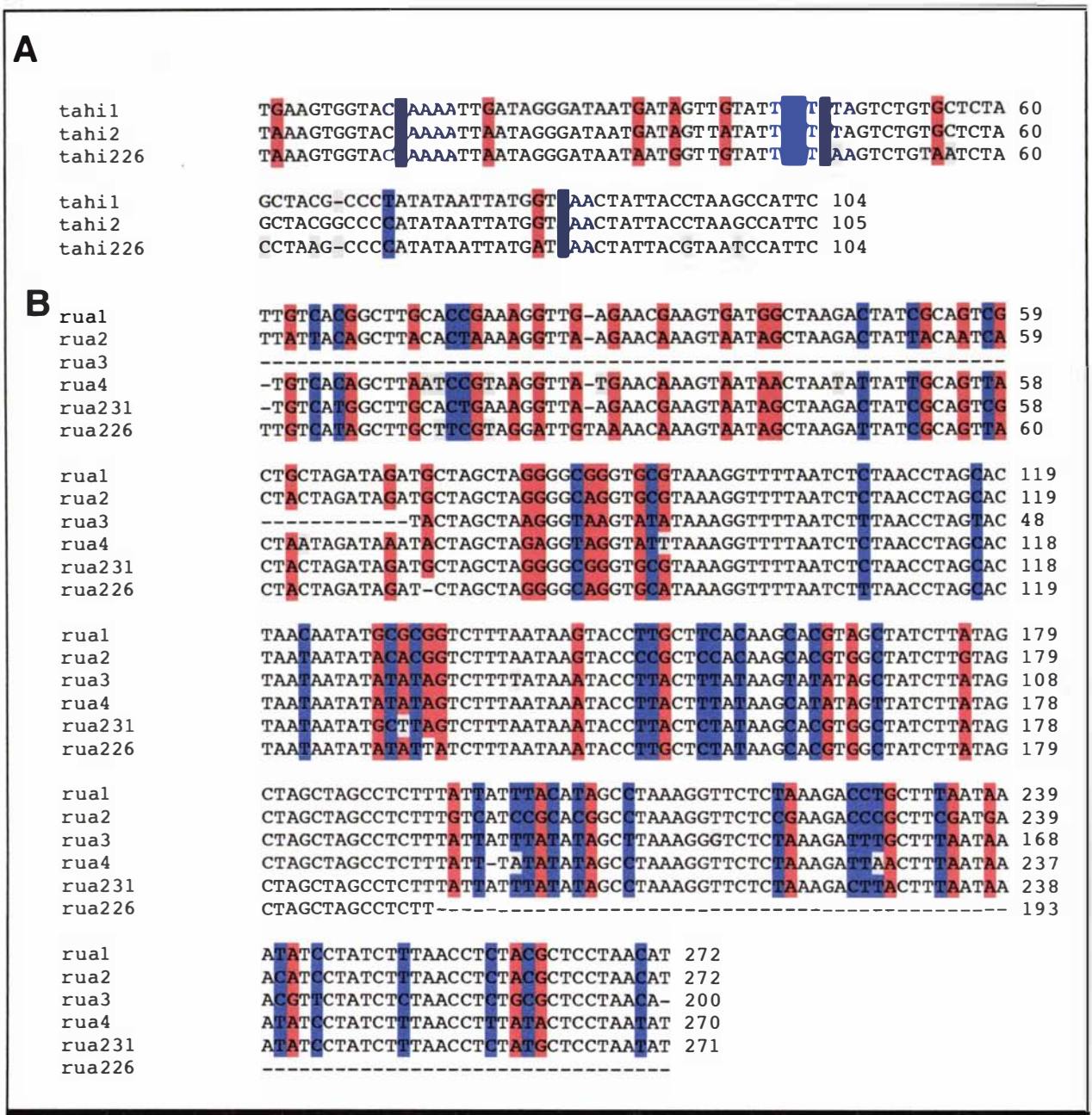


Figure 3.38 Alignment of the Tahi and Rua LTR sequences

(A) Tahi1 and Tahi2 are the two Tahi LTR sequences found in the retrotransposon platform between Lp19 *ltmK* and *pks*. Tahi226 was identified in a lambda clone λCY226 isolated with the Rua probe (amplified with primers lol4 and lol23). **(B)** Rua1 and Rua2 are the two Rua LTR sequences found in the retrotransposon platform between *ltmK* and *pks*. Rua3 is the LTR adjacent to *ltmG*. Rua4 was identified adjacent to *ltm* cluster 2 (Section 3.9.1). Rua231 and Rua226 are Rua LTR sequences identified within lambda clones λCY231 and λCY226, respectively that were isolated during a library screen with the Rua probe.

The red highlighting shows base differences of G and A, the blue highlighting shows base differences of C and T, while grey highlighting indicates other possible changes.

3 CpA (TpG) and 6 TpG from 12 transitions. The Rua LTR sequences had 59% transitions that were unequivocally considered to have a CpA (TpG) consensus, as there were 9 CpA and 13 TpG from 37 transitions. Twelve of the 15 transitions that were not the result of a CpA dinucleotide had a consensus of CpT. As found in *N. crassa* (Selker, 1990), *Podospora anseriana* (Graïa, et al., 2001) and *Leptosphaeria maculans* (Idnurm, Howlett, 2003) the sequence bias of transitions in the *N. lolii* retrotransposons was CpAs followed by CpTs.

3.9 Isolation of two additional *ltm* gene clusters

The presence of the retrotransposon platforms and the instability of cloned fragments of these flanking regions made it difficult to isolate additional sequences flanking the cluster of *ltmG, ltmM* and *ltmK*. Based on our knowledge of the genes required for paxilline biosynthesis in *P. paxilli* (Young, et al., 2001; McMillan, et al., 2003) and aflatrem biosynthesis in *A. flavus* (Zhang, et al., 2004), and the structural similarities of these compounds to lolitrem B, orthologues of *paxC*, *paxP* and *paxQ* were predicted to be present in *N. lolii*. Alignments of PaxC and AtmC polypeptide sequences were used to identify conserved regions to design degenerate primers for isolation of the *N. lolii* orthologue, *ltmC* (Appendix 5.1). A number of PCR amplification conditions with the four primer combinations were trialed but no evidence of an identifiable product of a *paxC* orthologue was amplified from *N. lolii* Lp19.

At this stage of the project, EST sequences from endophyte-infected ryegrass, which showed significant homology to *paxC* and *paxP*, were made available (by G. Spangenberg, Plant Biotechnology Centre, La Trobe University, Melbourne) to test the hypothesis that these EST sequences corresponded to *ltm* biosynthetic genes that were linked to the cluster containing *ltmG*, *ltmM* and *ltmK*. A summary of these sequences with the match to their *pax* homologue and the region of the gene they span is shown in Table 3.9 and Figure 3.39. Of interest were fragments with similarities to *paxC* and *paxP*, a prenyl transferase and cytochrome P450 monooxygenase respectively (Table 3.9). The EST sequences obtained did not align to the cluster containing *ltmG*, *ltmM* and *ltmK* suggesting they were unique. As the quality of the sequences was variable, the ESTs were aligned into contigs using the sequencher software to minimise errors for subsequent primer design. Primers were then designed to regions that were highly conserved to their *pax* homologues with a consideration on the placement of possible

Table 3.9 EST sequences of potential *ltm* genes

Gene	Sequence identification[1]	*pax* gene	aa range[2]	E value[3]	% identity[3]	Primers[4]	PCR size (bp)
ltmC	06nl1GsG12 06nl1FsB04 06nl1EsB02 06nl1EsG01	*paxC*	35-174	2e-26	60%	lol189/190	360
	06nl1BsC09	*paxC*	172-229	1e-07	49%	lol190/193	644
ltmP	06nl1BsA06 06nl1CsG09 RJ-J15	*paxP*	47-304	1e-40	36%	lol191/192	374
	07nl1AsA11 RJ-G13	*paxP*	307-414	4e-19	44%	lol192/195	603
ltmJ	06nl1DsF08 RJ-N17	*paxP*[5]	414-495	7e-05	38%	lol205/206	242

[1]The EST identification number of the sequences provided by G. Spangenberg (Plant Biotechnology Centre, La Trobe University, Melbourne) and R. Johnson (AgResearch, Palmerston North).

[2]The amino acid range of the match of the EST sequences to the *pax* homologue.

[3]The E-value of the BLASTX match or the % identity to the *pax* homologue

[4]Primers used to amplify a genomic Lp19 fragment

[5]The best match was to a P450 monooxygenase from *Asperillus nidulans* (AN1598), at 3e-09

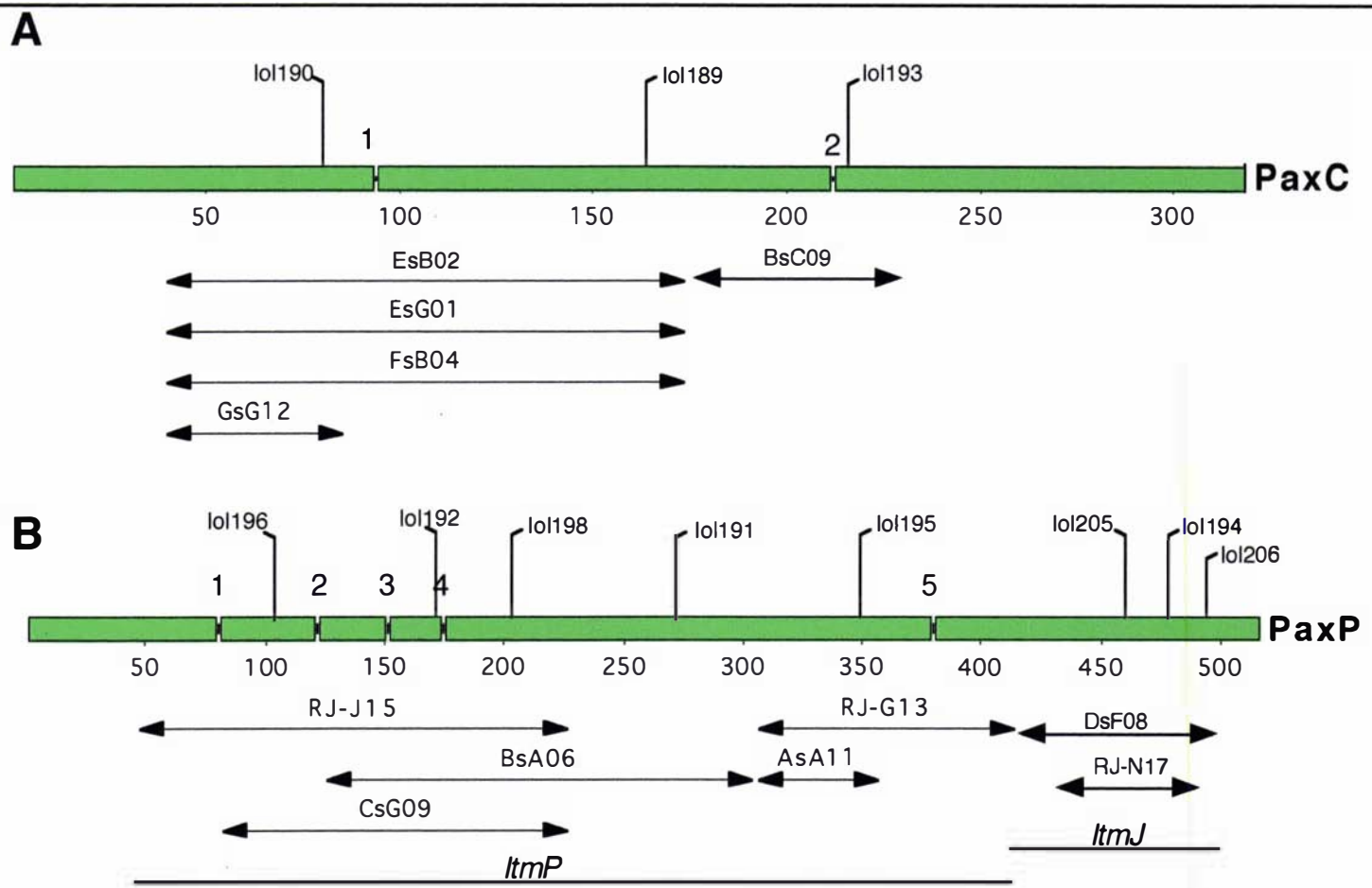


Figure 3.39 Schematic diagram of PaxC and PaxP showing placement of the EST sequences

Schematic diagram of (**A**) PaxC and (**B**) PaxP, showing the relative placement of the EST sequences. The polypeptide sequence is represented as green blocks with the size indicated in amino acid residues underneath. The intron placements are numbered above the polypeptide. The primers used for PCR amplification are positioned above the region used for primer design. The EST sequences that are part of the *ltmP* or the *ltmJ* gene are shown as lines below the EST positions. The EST identification numbers (Table 3.9) have been reduced to the last five numbers.

conserved introns between the *ltm* and *pax* genes. Additional EST sequences were made available (by R. Johnson, AgResearch, Palmerston North) from a suppression subtraction hybridisation technique, where three transcripts, RJ-J15, RJ-G13 and RJ-N17, were identified as genes up-regulated in *N. lolii* Lp19-infected perennial ryegrass when compared to uninfected perennial ryegrass. These three EST sequences were also included in the sequence analysis data (Table 3.9; Fig. 3.39).

The contig of EST fragments, 06nl1GsG12, 06nl1FsB04, 06nl1EsB02 and 06nl1EsG01, was designated *ltmC* based on their significant match to PaxC using the BLASTX algorithm (Table 3.9). A fragment of 360 bp, the same size as that of the cDNA, was amplified from Lp19 genomic DNA using primers lol189 and lol190. This confirmed that the EST fragments were of fungal origin. PCR amplification was used to join the *ltmC* contig to 06nl1BsC09 a potential *ltmC* fragment (Table 3.9; Fig. 3.39). Amplification from *N. lolii* Lp19 genomic DNA with primers lol190 and lol193 gave a 644-bp product that was 77-bp larger than the predicted transcript size, indicating the presence of an intron in this gene (Fig. 3.39). The PCR product generated from Lp19 genomic DNA with primers lol190 and lol193 was sequenced and compared to the EST sequences for confirmation of the intron.

The contig of EST fragments, 06nl1BsA06 and 06nl1CsG09 (Fig. 3.39), had a significant match to *paxP*, spanning amino acid residues 84 through 304 (Table 3.9). Primers, lol191 and lol192, designed to this sequence amplified a 374-bp fragment from *N. lolii* Lp19 genomic DNA. This confirmed these EST sequences were of fungal origin. The size of the genomic fragment was 60 bp larger than the predicted cDNA product indicating that this gene contains at least one intron.

EST sequences with the BLASTX matches to *paxP* aligned into three independent contigs (Table 3.9; Fig. 3.39). Contig 1 contained EST sequences 06nl1BsA06, 06nl1CsG09 and subsequently RJ-J15, contig 2 contained EST sequences 07nl1AsA11 and RJ-G13, and contig 3 contained EST sequences 06nl1DsF08 and RJ-N17. PCR was performed to test whether these three contigs were part of a single cytochrome P450 monooxygenase gene or were in fact multiple genes. Amplification of *N. lolii* Lp19 genomic DNA with primers lol192 and lol195 linked contigs 1 and 2 and therefore these two contigs are a part of the same fungal cytochrome P450 monooxygenase gene subsequently named *ltmP*. The PCR fragment generated from *N.*

lolii Lp19 genomic DNA with primers lol192 and lol195 was sequenced and compared to the EST data for confirmation of the intron. Contig 3 contained the primer-binding site for primer lol194 and this primer would not amplify a PCR product from Lp19 genomic DNA when paired with primer lol192. This contig was therefore considered an independent cytochrome P450 monooxygenase fragment and was subsequently named *ltmJ*. Primers, lol205 and lol206, were designed to the contig sequence of *ltmJ*. These primers amplified a 242-bp fragment from Lp19 genomic DNA and confirmed that *ltmJ* was of fungal origin.

Hybridisation of an *ltmC* fragment to genomic digests of the endophyte strains Lp19, Fl1, and E8 shows that this fragment is only present in the lolitrem-producing strains, Lp19 and Fl1 (Fig. 3.40A). The *ltmC* hybridisation pattern of each digest was identical between the Lp19 and Fl1 strains, but the hybridisation signal in Lp19 was twice that of Fl1 despite identical DNA loadings.

Two *ltmP* fragments were used as hybridisation probes to digested genomic DNA of strains *N. lolii* Lp19, *E. festucae* Fl1 and *E. typhina* E8. Each fragment was only detected in the Lp19 and Fl1 strains (Fig. 3.40). The *ltmP* fragment amplified with primers lol191 and lol192 gave different Lp19 and Fl1 banding patterns with each restriction enzyme (Fig. 3.40C). The *ltmP* fragment amplified with primers lol196 and lol198 spans both a *Hin*dIII and *Sst*I restriction enzyme site (Fig. 3.41). Hybridisation with this fragment resulted in two *Hin*dIII and *Sst*I hybridising bands. The 0.5-kb *Hin*dIII and 9 kb *Sst*I fragments in each digest were the same size in the Lp19 and Fl1 digested DNA (Fig. 3.40B). The 9-kb *Sst*I fragment hybridised to both the *ltmP* fragment, amplified with primers lol196 and lol198, and to the *ltmC* fragment. These data show that the *ltmC* and *ltmP* genes are linked.

The *ltmJ* fragment hybridised to the lolitrem-producing strains *N. lolii* Lp19 and *E. festucae* Fl1 (Fig. 3.40). This fragment hybridised to a ~18 kb Lp19 *Sst*I fragment, a band of the same size as seen with the *ltmP* probes suggesting linkage of *ltmJ* to *ltmP*. The presence of the three EST fragments, *ltmC*, *ltmP* and *ltmJ*, correlated with strains known to produce indole-diterpenes. None of the fragments hybridised to genomic digests of E8, a lolitrem non-producing strain. This pattern of hybridisation was used to identify the previous *ltm* cluster containing *ltmG*, *ltmM* and *ltmK*, and therefore, complete sequence surrounding the genes *ltmC*, *ltmP* and *ltmJ* was obtained.

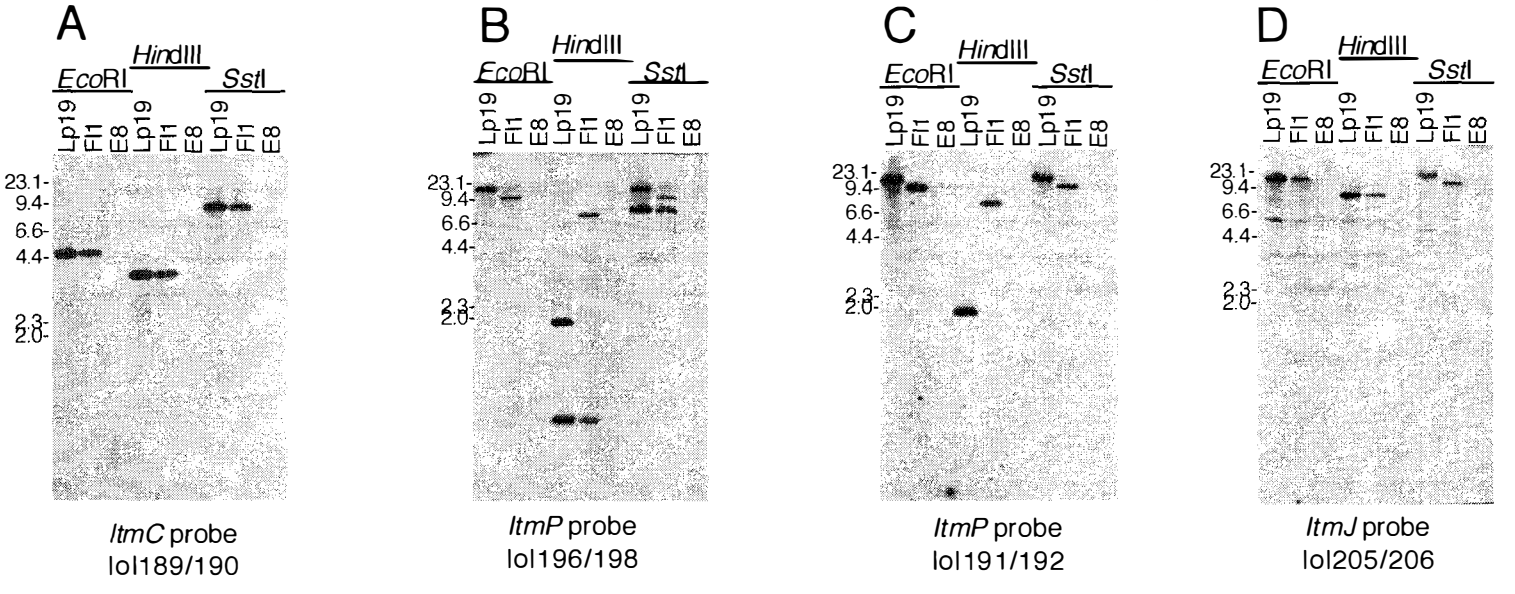


Figure 3.40 Southern analysis of putative *ltm* genes

Autoradiographs of Southern hybridisation of ^{32}P-labelled fragments of (**A**) *ltmC*, amplified with primers lol189 and lol190, (**B**) *ltmP*, amplified with primers lol196 and lol198, (**C**) *ltmP*, amplified with primers lol191 and lol192, and (**D**) *ltmJ*, amplified with primers lol205 and lol206, to *Eco*RI, *Hin*dIII and *Sst*I restriction enzyme digests of *N. lolii* Lp19, *E. festucae* Fl1 and *E. typhina* E8. The size standards are in kb.

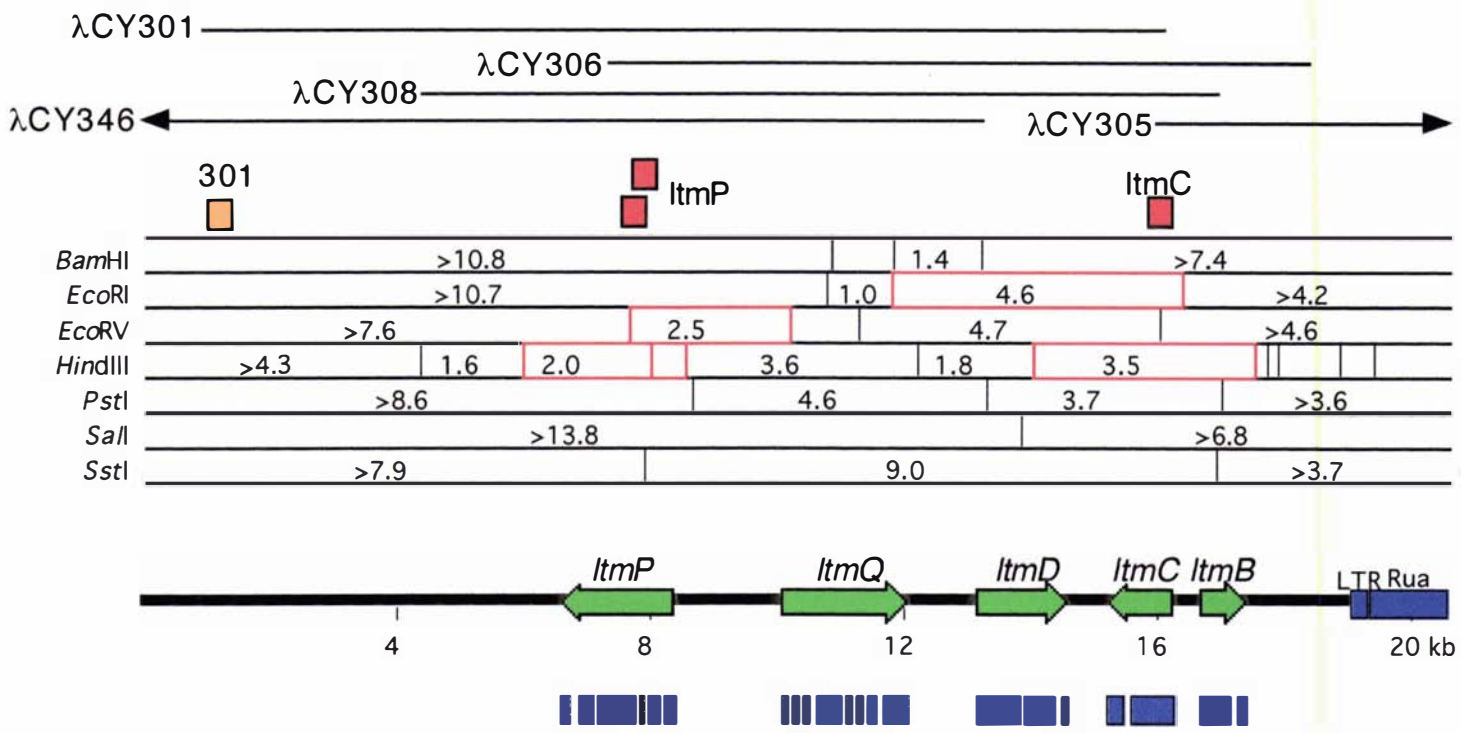


Figure 3.41 A physical map of the *Itm* cluster 2

A physical and genetic map pf the Lp19 *Itm2* locus. The five *Itm* genes are shown as green arrows, the exons of the *Itm* genes are blue boxes under the gene. Selected lambda clones isolated with the *ItmC* and *ItmP* probes are shown as black lines. The fragments used as probes to isolate the lambda clones and for the Southern analysis in Fig. 3.40 are shown as red boxes above the restriction enzyme map. The fragments initially isolated by IPCR are boxed in red. The fragment used as a probe in Fig. 3.43 is shown as an orange box.

3.9.1 *ltm* cluster 2

Initially the complete *ltmC* and *ltmP* genes were sequenced from Lp19 using fragments generated by IPCR with the restriction enzymes *Eco*RI, *Eco*RV and *Hin*dIII (Fig. 3.41). The complete *ltmC* gene was amplified using IPCR with Lp19 *Hin*dIII digested then self-ligated genomic DNA and primers, lol202 and lol203, that were designed to the previously obtained *ltmC* sequence. The sequence was extended using IPCR with Lp19 *Eco*RI-digested then self-ligated genomic DNA and primers lol213 and lol209. The sequence of the complete *ltmP* gene was generated using IPCR with Lp19 *Hin*dIII-digested then self-ligated genomic DNA using two primer sets of lol198 and lol199, and lol210 and lol211. The sequence was extended further by IPCR using Lp19 *Eco*RV-digested then self-ligated genomic DNA with primers lol192 and lol222. Each IPCR fragment was cloned into pGEM-T easy (Promega) and sequenced with primers that were *ltmC* or *ltmP* sequence-specific or with primers Sp6 and T7.

To complete the sequence analysis of this region, an *N. lolii* Lp19 λGEM-12 genomic library was screened with *ltmC* and *ltmP* (Section 2.5). Using the *ltmC* fragment amplified with lol189 and lol190 the library screen resulted in the isolation of 24 positive clones (Table 3.10) from ~80,000 plaques screened. The same Lp19 λGEM-12 genomic library filters were screened with the *ltmP* fragment, amplified with lol191 and lol192, and resulted in the isolation of 25 positive clones of which 18 also hybridised with *ltmC* (Table 3.10). This result confirmed the Southern analysis showing that *ltmC* and *ltmP* are contained on the same 9 kb *Sst*I fragment (Fig. 3.40; 3.41). The average insert size of the 35 lambda clones was approximately 13 kb. DNA isolated from the lambda clones was digested with the restriction enzymes *Bam*HI, *Eco*RI, *Hin*dIII and *Sst*I, then hybridised with the *ltmC* and *ltmP* fragments to determine clones for sequencing. Lambda clones of interest were sequenced with primers Sp6 and T7 that anneal to the lambda arms and then with sequence-specific primers. To facilitate sequencing, fragments from some lambda clones were cloned into the pUC118 vector and sequenced with the forward and reverse primers (Section 2.5). A physical map of the overlapping lambda clones (Fig. 3.41) was determined based on DNA sequence analysis (Appendix 5.3) and the hybridisation data from both the lambda clones and the genomic DNA (Fig. 3.40). Data generated from the physical map of the lambda clones showed that the following clones, λCY300, λCY307, λCY312, λCY313, λCY315, λCY316, λCY319 and λCY350, contained sequences that were inconsistent with the

Table 3.10 A list of the lambda clones identified from screening the Lp19 λGEM-12 genomic library with the *ltmC, ltmP* and *ltmJ* fragments

Lambda clone	*ltmC* probe[1]	*ltmP* probe[1]	*ltmJ* probe[1]	Identical to[2]	comments[3]
λCY300	*ltmC*	*ltmP*			Rearranged T7 end
λCY301	*ltmC*	*ltmP*		304, 320	Sequence
λCY302	*ltmC*	*ltmP*			DNA contains more than one clone
λCY303	*ltmC*	*ltmP*		322	Sequence
λCY304	*ltmC*	*ltmP*		301, 320	Sequence
λCY305	*ltmC*				Rearranged Sp6 end. Sequence
λCY306	*ltmC*	*ltmP*		318, 321	Sequence
λCY307	*ltmC*	*ltmP*			Rearranged T7 end. Sequence
λCY308	*ltmC*	*ltmP*			Sequence
λCY309	*ltmC*	*ltmP*			
λCY310	*ltmC*	*ltmP*			
λCY311	*ltmC*	*ltmP*			DNA contains more than one clone
λCY312	*ltmC*				Rearranged T7 end
λCY313	*ltmC*			316	Rearranged Sp6 end
λCY314	*ltmC*	*ltmP*			
λCY315	*ltmC*				Rearranged end
λCY316	*ltmC*			313	Rearranged end
λCY317	*ltmC*	*ltmP*			Sequence
λCY318	*ltmC*	*ltmP*		306, 321	
λCY319	*ltmC*				
λCY320	*ltmC*	*ltmP*		301, 304	Sequence
λCY321	*ltmC*	*ltmP*		306, 318	
λCY322	*ltmC*	*ltmP*		303	
λCY323	*ltmC*	*ltmP*			Never isolated DNA
λCY324			*ltmJ*	344	Sequenced
λCY325			*ltmJ*	338	Rearranged T7 end. Sequenced *chsV*.
λCY326			*ltmJ*	329, 333, 334, 335, 339, 343	Rearranged T7 end. Sequence.
λCY327			*ltmJ*		Never isolated DNA
λCY328			*ltmJ*		Not *ltmJ* positive after round 3
λCY329			*ltmJ*	326, 333, 334, 335, 339, 343	Rearranged T7 end. Sequence.
λCY330			*ltmJ*		
λCY331			*ltmJ*		Never isolated DNA
λCY332			*ltmJ*		Rearranged Sp6 end. Sequence.
λCY333			*ltmJ*	326, 329, 334, 335, 339, 343	
λCY334			*ltmJ*	326, 329, 333, 335, 339, 343	
λCY335			*ltmJ*	326, 329, 333, 334, 339, 343	
λCY336			*ltmJ*		DNA contains more than one clone
λCY337			*ltmJ*		
λCY338			*ltmJ*	325	Rearranged T7 end. Sequenced *chsV*.
λCY339			*ltmJ*	326, 329, 333, 334, 335, 343	
λCY340			*ltmJ*		Not *ltmJ* positive after round 3
λCY341			*ltmJ*		Never isolated DNA
λCY342			*ltmJ*		
λCY343			*ltmJ*	326, 329, 333, 334, 335, 339	
λCY344			*ltmJ*	324	
λCY345			*ltmJ*		Never isolated DNA
λCY346		*ltmP*			Sequence
λCY347		*ltmP*			
λCY348		*ltmP*			
λCY349		*ltmP*			
λCY350		*ltmP*			Strange digestion pattern
λCY352		*ltmP*			
λCY353		*ltmP*			

[1] *ltmC* indicates positive for the *ltmC* fragment amplified with primers lol189 and lol190
ltmP indicates positive for the *ltmP* fragment amplified with primers lol191 and lol192
ltmJ indicates positive for the *ltmJ* fragment amplified with primers lol205 and lol206

[2] Lists the clones that are identical based on mapping and sequence data

[3] The comments indicate what end of the lambda clone is rearranged (when known)
Sequence means that the DNA was used as a sequencing template and part of the sequence aligned within a cluster

Lp19 genomic map, and therefore, these sequences were not analysed further. The mapped genomic region surrounding the *ltmC* gene was underrepresented by the lambda clones, which was a similar finding to that of the *ltmG* library screen (Section 3.3 and 3.7). A list of the lambda clones identified from the Lp19 λGEM-12 library is summarised in Table 3.10

Nucleotide sequence generated from sequencing lambda clones isolated from the hybridisation with *ltmC* and *ltmP* covered 23.8 kb and confirmed the linkage of *ltmC* to *ltmP*. Sequence analysis of this region using BLAST algorithms identified three additional genes with significant similarities to *pax* genes from *P. paxilli*, and therefore, these genes formed a second gene cluster called *ltm* cluster 2 (Table 3.11; Fig. 3.41). These genes included *ltmC*, *ltmP*, *ltmQ*, *ltmD* and *ltmB*, orthologues of *paxC* (prenyl transferase), *paxP* and *paxQ* (cytochrome P450 monooxygenases), *paxD* (prenyl transferase, dimethylallyl tryptophan synthase-like) and *paxB*, (gene of unknown function), respectively (Table 3.11; 3.12; Fig. 3.41). The *ltmJ* gene was not contained within this sequenced region. The individual sequence analysis of the five genes contained in *ltm* cluster 2 is explained below and in Tables 3.11 and 3.12.

The nucleotide sequence analysis of the complete *ltmC* gene, initially identified from EST sequences (Section 3.9), showed that it contains one intron (Fig. 3.42) and encodes a polypeptide of 345 amino acids. LtmC is classified as a prenyl transferase as it contains the five conserved domains found in other prenyl transferases (Chen, et al., 1994) (Table 3.11). FastA analysis showed that LtmC was more similar to AtmC from *A. flavus* than PaxC from *P. paxilli* (Table 3.12). The single intron found in *ltmC* was conserved with placement and phase with the second of the two introns found in *P. paxilli paxC* (Young, et al., 2001) and *A. flavus atmC* (Zhang, et al., 2004) (Appendix 5.1).

Sequence analysis of the complete *ltmP* gene, initially identified from EST sequences (Section 3.9), showed that it contained five introns (Fig. 3.42) and encodes a polypeptide of 498 amino acids (Table 3.11). LtmP is classified as a cytochrome P450 monooxygenase based on database matches. The placement and phase of four introns, 1, 2, 3 and 4, are conserved with the *paxP* introns, 1, 3, 4 and 5 (Young, et al., 2001); and three introns 1, 3 and 4, are conserved with *ltmK* introns, 1, 4 and 7 (Appendix 5.1). LtmP is more similar to PaxP than to PaxQ or LtmK (Tables 3.11; 3.12).

Table 3.11 The *ltm* genes from clusters 2 and 3, intron analysis and comparisons to database sequences

Gene	Putative function	Cluster	size (aa)	kDa	Intron				Top Database match[1]	Species	E value	Reference
					No.	phase	size	5'...3' Splice sites				
ltmB	Unknown	2	221	24.4	1	2	76	GTAAGT...CAG	paxB[2]	*P. paxilli*	2e-59	
ltmC	Prenyl transferase	2	345	39	1	2	77	GTATGT...TAG	paxC	*P. paxilli*	1e-59	Young et al 2001
ltmD	Prenyl transferase (*dmaW*-like)	2	439	49	1	2	74	GTAAGA...CAG	AN8514	*A. nidulans*	1e-46	
					1	2	102	GTAAGT...TAG				
ltmQ	P450 monooxygenase	2	537	61.4	1	0	59	GTTTGA...AAG	paxQ	*P. paxilli*	1e-105	Young et al 2001
					2	0	61	GTTTGT...TAG				
					3	0	59	GTAAGT...CAG				
					4	2	60	GTAAGC...TAG				
					5	0	52	GTATGG...TAG				
					6	0	53	GTATAT...TAG				
					7	1	56	GTATAA...CAG				
ltmP	P450 monooxygenase	2	498	57	1	0	59	GTGTTC...CAG	paxP	*P. paxilli*	1e-102	Young et al 2001
					2	1	49	GTAAGT...CAG				
					3	1	60	GTATGT...GAG				
					4	1	57	GTAAGG...AAG				
					5	0	107	GTATGT...TAG				
ltmJ	P450 monooxygenase	3	525	60.4	1	0	63	GTGAGC...CAG	AN1598	*A. nidulans*	4e-81	
					2	0	67	GTGAGG...TAG				
					3	0	49	GTAAGG...AAG				
					4	1	69	GTATGT...TAG				
					5	1	61	GTAAGA...CAG				
					6	1	58	GTAAGT...TAG				
ltmE	Prenyl transferase fusion (*paxC-paxD*)	3	788	87.5	1	2	67	GTACGT...AAG	paxC[3]	*P. paxilli*	3e-60	Young et al 2001
					2	2	56	GTGAGT...CAG	AN8514[3]	*A. nidulans*	7e-56	
					3	2	54	GTACGT...TAG				
chsV	Chitin synthase (class V)		1861	206	1		82	GTAGGT...TAG	AAF04279	*Blumeria graminis*	0.0	Zhang et al 2000
					2		69	GTAAGT...CAG				

[1] If the top match was to a *pax* gene, the gene name was used otherwise the gene is indicated by the accession or gene number.

[2] The top match was determined using the tBLASTn algorithm as the *paxB* gene is currently not annotated

[3] This gene is a hybrid containing a *paxC*-like domain and a *paxD*-domain

Table 3.12 Sequence identity of the *ltm* genes to their *pax* and *atm* homologues

gene	Homologue	*Species*	% Identity	E-value	Analysis programme
ltmG	*ggs1*	*P. paxilli*	54.1	1e-90	FASTA/BLASTP
	paxG	*P. paxilli*	52.6		FASTA
	paxC	*P. paxilli*	31.5		FASTA
	paxD	*P. paxilli*	22.2		FASTA
	atmG	*A. flavus*	59.4	e-101	FASTA/BLASTP
	atmC	*A. flavus*	30.1		FASTA
	ltmC	*N. lolii*	28.4		FASTA
	ltmD	*N. lolii*	28.4		FASTA
	ltmE	*N. lolii*	31.5		FASTA
ltmM	*paxM*	*P. paxilli*	41	7e-96	BLASTP
	atmM	*A. flavus*	42.2	e-100	BLASTP
ltmK	*paxP*	*P. paxilli*	31.3	7e-63	FASTA/BLASTP
	paxQ	*P. paxilli*	23.4		FASTA
	ltmJ	*N. lolii*	36.8		FASTA
	ltmP	*N. lolii*	28.6		FASTA
	ltmQ	*N. lolii*	25.3		FASTA
ltmB	*paxB*	*P. paxilli*	53.8	2e-59	tBLASTn
	FG04594	*Fusarium graminearum*		4e-46	BLASTP
ltmC	*paxC*	*P. paxilli*	43.3	1e-59	FASTA/BLASTP
	paxG	*P. paxilli*	28.4		FASTA
	atmC	*A. flavus*	47.7	2e-68	FASTA/BLASTP
	atmG	*A. flavus*	28.1		FASTA
	ltmE	*N. lolii*	55.8		FASTA
	ltmG	*N. lolii*	28.4		FASTA
ltmD	*paxD*	*P. paxilli*	24.2		FASTA
	ltmE	*N. lolii*	37.1		FASTA
	ltmG	*N. lolii*	24.2		FASTA
	dmaW	*Neotyphodium* LpTG-2	22.5		FASTA
	AN8514	*Aspergillus nidulans*		1e-46	BLASTP
ltmP	*paxP*	*P. paxilli*	41.3	e-102	FASTA/BLASTP
	paxQ	*P. paxilli*	24.4		FASTA
	ltmJ	*N. lolii*	25		FASTA
	ltmK	*N. lolii*	29.2		FASTA
	ltmQ	*N. lolii*	24.5		FASTA
ltmQ	*paxQ*	*P. paxilli*	38.1	e-105	FASTA/BLASTP
	paxP	*P. paxilli*	28.7		FASTA
	ltmJ	*N. lolii*	22.2		FASTA
	ltmK	*N. lolii*	25.3		FASTA
	ltmP	*N. lolii*	24.1		FASTA
ltmJ	*paxP*	*P. paxilli*	29.2	1e-49	FASTA/BLASTP
	paxQ	*P. paxilli*	23.7		FASTA
	ltmK	*N. lolii*	36.8		FASTA
	ltmP	*N. lolii*	25		FASTA
	ltmQ	*N. lolii*	21.9		FASTA
	AN1598	*A. nidulans*		4e-81	BLASTP
ltmE	*paxC*	*P. paxilli*	43.1	3e-60	FASTA/BLASTP
	atmC	*A. flavus*	49.5	7e-71	FASTA/BLASTP
	ltmC	*N. lolii*	55.8		FASTA
	ltmD	*N. lolii*	37.1		FASTA
	ltmG	*N. lolii*	32.8		FASTA
	AN8514	*A. nidulans*		7e-56	BLASTP

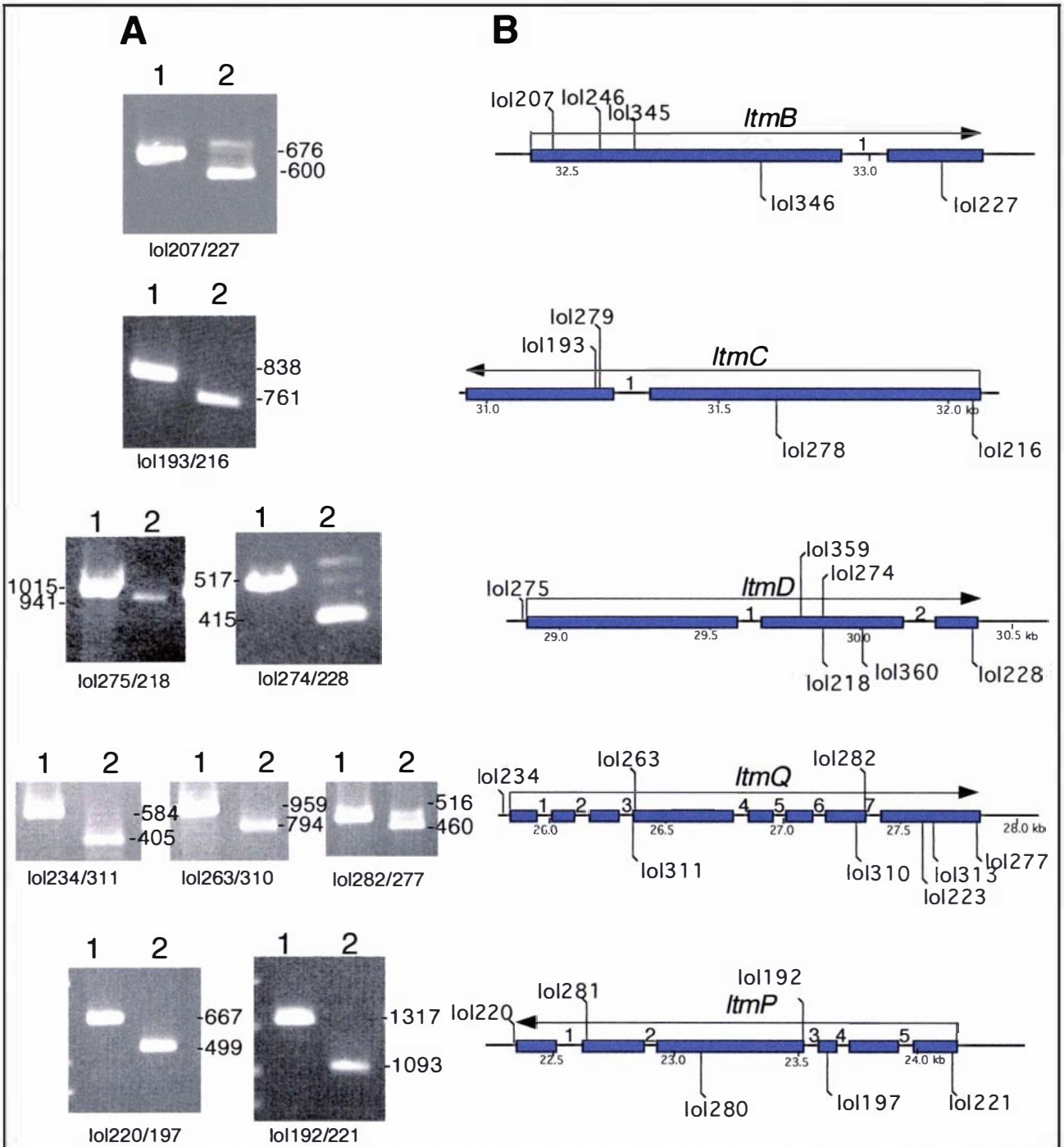


Figure 3.42 Gene structures of the *ltm* cluster 2 genes

Gene structures of the *ltm* cluster 2 genes. (**A**) RT-PCR confirming placement of introns. Lane (1) Lp19 genomic DNA; (2) cDNA from Lp19-infected Nui perennial ryegrass. The RT-PCR products were purified and sequenced with gene-specific primers. Primer combinations used are indicated under each gel picture. The sizes of the PCR products are indicated in bp. (**B**) Gene structures showing introns and placement of primers used for RT-PCR analysis. The exons are represented by blue boxes and the primers are positioned above or below the sequence for forward or reverse direction respectively. The introns are numbered between the exons. The direction of transcription of the gene is indicated by the arrow above the exons. The approximate position of the gene within the cluster (see Figure 3.50) is indicated in kb.

RT-PCR analysis was performed with cDNA that was reverse transcribed using mRNA purified from Lp19-infected Nui perennial ryegrass. The PCR reaction conditions were: 5 ng genomic DNA or 5 μL of diluted cDNA, 1x *Taq* polymerase buffer (Roche), 50 μM each dNTP, 200 nM each primer, 0.5 U *Taq* polymerase, in a reaction volume of 25 μL. The PCR amplification conditions were as follows: 1 cycle of 94^{o}C for 2 min; 35 cycles of 94^{o}C for 15 s, 60^{o}C for 30 s, 72^{o}C for 1 min; followed by one cycle of 72^{o}C for 10 min. The extension time was adjusted to 1 min/kb for larger fragments.

Adjacent to *ltmP* is *ltmQ*, a cytochrome P450 monooxygenase gene (Fig. 3.41). The best database match to *ltmQ* is that of *paxQ* from *P. paxilli* (Tables 3.11; 3.12) and FastA analysis confirmed that LtmQ is more similar to PaxQ than to PaxP (Table 3.12). The *ltmQ* gene contains 7 introns (Fig. 3.42) of which six introns, 2 to 7, are conserved in placement and phase with *paxQ* introns 2, 3, 5, 6, 7 and 8 (Appendix 5.1). It appears from ClustalW alignments of the *ltm* and *pax* P450 monooxygenases (Appendix 5.1) that *ltmQ* has lost a conserved intron that is present in the remaining five sequences shown in the alignment, corresponding to a position between the current *ltmQ* introns 3 and 4. The intron boundaries of *ltmQ* were confirmed by sequence comparison of RT-PCR products amplified using cDNA from endophyte-infected plant material and gene-specific primers to the Lp19 genomic region (Fig 3.42; Appendix 5.3).

The *ltmB* gene product has no predicted function and had a best BLASTP match to an uncharacterised gene from *F. graminearum* FG04594 (accession number EAA72208). Using the tBLASTN algorithm against the public databases, the best match was to *paxB*, a gene recently identified within the *P. paxilli pax* cluster, but as yet not publicly annotated (Monahan and Scott, unpublished). The *ltmB* gene has one intron (Fig. 3.42) that is conserved in placement and phase with the *paxB* gene from *P. paxilli* (Appendix 5.1).

The *ltmD* gene had a best BLASTP match to *A. nidulans* AN8514 (accession number EAA66858). As the publicly available *paxD* sequence was incomplete, FastA analysis was used to compare LtmD to DmaW from Lp1 a *Neotyphodium* LpTG-2 and PaxD from *P. paxilli* (Tables 3.11; 3.12). This data showed that LtmD is more similar to PaxD than to the DmaW from *Neotyphodium* LpTG-2. The *ltmD* gene contains two introns of which the placement and phase of intron 2 are conserved with *paxD* (Appendix 5.1).

The predicted introns of the five genes contained in *ltm* cluster 2 were confirmed by sequence comparison of cDNA sequences, generated by RT-PCR using cDNA from endophyte-infected plant material, with the genomic sequence (Fig. 3.42; Appendix 5.3). A summary of the intron numbers, intron splice sites and predicted molecular mass in kDa of each gene product is shown in Table 3.11.

Flanking *ltmB* is a Rua long terminal repeat (Rua4 in Fig. 3.38) and degenerate retrotransposon sequence (Fig. 3.41). Upstream from *ltmP* is an AT-rich region that was devoid of obvious open reading frames and no genes were evident from sequence analysis using BLAST searches. Southern analysis with a fragment from this region to *Eco*RI-, *Hin*dIII- and *Sst*I-digested DNA showed that this sequence is present in the lolitrem-producing strains *N. lolii* Lp19 and *E. festucae* Fl1, but is absent from the non-producer *E. typhina* E8 (Fig. 3.43). Based on Southern analysis there are predicted to be ~3 - 5 copies of this sequence contained within the Lp19 and Fl1 genomes (Fig. 3.43). The presence of AT-rich sequences adjacent to *ltmB* and *ltmP* suggested that no additional genes are present at this locus thereby defining the boundaries of *ltm* cluster 2.

3.9.2 *ltm* Cluster 3

The DNA sequence surrounding *ltmJ* that encodes a cytochrome P450 monooxygenase initially identified from EST sequences (Section 3.9), was isolated from the *N. lolii* Lp19 λGEM-12 genomic library hybridised with the *ltmJ* fragment, amplified with primers lol205 and lol206. This hybridisation resulted in the isolation of 22 positive clones (Table 3.10). Fifteen clones were digested with *Hin*dIII or *Bam*HI and hybridised with the *ltmJ* fragment to determine clones of interest (Table 3.10). Comparison of the restriction enzyme digests and sequencing of these clones, with primers SP6 and T7 that anneal to the lambda arms, showed that only two identical clones, λCY324 and λCY344, had the correct genomic arrangement based on Southern and PCR analysis. Other lambda clones were rearranged and/or contained unrelated sequences (Appendix 5.3).

From sequence analysis, λCY346 and λCY324, identified from *ltmP* and *ltmJ* hybridisations respectively, were shown to overlap, thus linking *ltm* clusters 2 and 3 with a 16-kb AT-rich region separating them. Sequence analysis of this AT-rich region, using the BLASTX analysis of this sequence, failed to identify any evidence of potential genes. The strong AT bias of this sequence introduces numerous stop codons strongly suggesting it is non-coding. Additional sequence flanking the left-hand side of λCY324 was extended by IPCR using *Cla*I-, *Xba*I- or *Hin*dIII-digested then self-ligated Lp19 genomic DNA and sequence-specific primers (Fig. 3.42). Analysis of the *ltm* cluster 3 sequence identified two genes, *ltmJ* and *ltmE*, encoding a cytochrome P450

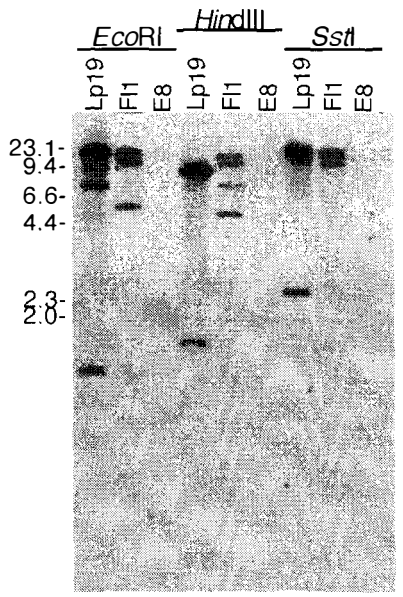


Figure 3.43 Southern analysis of the AT-rich region flanking *ltmP*

Autoradiographs of Southern hybridisation of a ^{32}P-labelled 301 fragment, amplified with primers lol253 and lol254 to *Eco*RI, *Hin*dIII and *Sst*I restriction enzyme digests of *N. lolii* Lp19, *E. festucae* Fl1 and *E. typhina* E8. The size standards are in kb. The position of the fragment used as a probe is shown in Fig. 3.41.

monooxygenase (*ltmJ*), and a multifunctional enzyme containing two prenyl-transferase domains, an LtmC type with a dimethylallyl tryptophan synthase LtmD type (Fig. 3.44).

The complete *ltmJ* gene was contained on λCY324 (Fig. 3.44). Sequence analysis of *ltmJ* revealed the presence of six introns (Fig. 3.45) of which all are conserved with the introns 1, 2, 3, 4, 5 and 7 from *ltmK* located in *ltm* cluster 1 (Section 3.4.2). LtmJ has a best BLASTP match to an *A. nidulans* AN1598 sequence (Tables 3.11; 3.12). Of the four *N. lolii* cytochrome P450 monooxygenase identified, LtmJ is most similar to LtmK followed by LtmP then LtmQ (Table 3.12).

The complete *ltmE* gene has significant BLASTP matches to both *P. paxilli* PaxC, and to the *A. nidulans* gene, AN8514 (Table 3.12). FastA analysis shows that LtmE is 55.8% identical to LtmC and 37.1% identical to LtmD. The *ltmE* gene contains 3 introns (Fig. 3.45) of which intron 1 from the *ltmC*-like domain is conserved with the placement and phase of *ltmC* intron, whereas intron 3 from the *ltmD*-like domain is conserved with the second *ltmD* intron.

The library screen completed with the *ltmJ* probe isolated two identical clones λCY325 and λCY338 that were rearranged at the T7 end. These clones contained a sequence with strong similarity to class V chitin synthase gene, *chsV*. The complete gene was contained within the clones and sequenced. The gene is approximately 5.7 kb long and has two introns (Fig. 3.45) that are conserved with placement and phase to those found in other fungal *chsV* genes. The sequence of the *chsV* gene is highly conserved with a significant BLASTP match to *Blumeria graminis* (accession number AAF04279) (Table 3.11). However, *chsV* is not part of *ltm* cluster 3.

The introns of *ltmJ* and *ltmE* from cluster 3 and the *chsV* from λCY325 and λCY338 were confirmed by RT-PCR using cDNA from endophyte-infected plant material (Fig. 3.45). The intron number, intron splice sites and predicted mass in kDa of each gene product are summarised in Table 3.11.

3.9.3 Expression profiles of the 10 *ltm* genes

The expression profiles of the 10 *ltm* genes, the *chsV* from λCY325 and the *pks* adjacent to cluster 1 were characterised *in planta* and in culture. Previous data showed that the

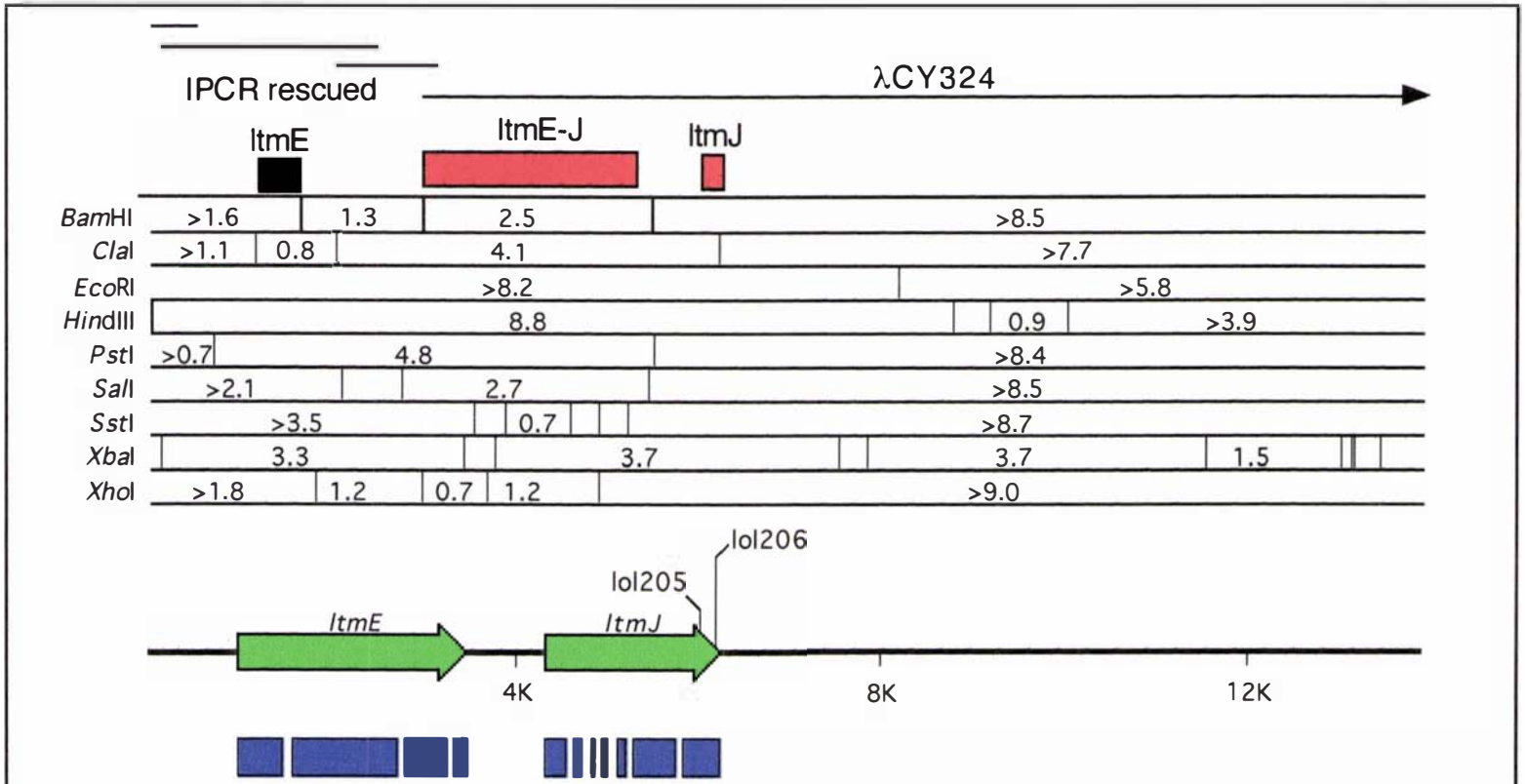


Figure 3.44 A physical map of *ltm* cluster 3

A physical and genetic map of the Lp19 *ltm* cluster 3 locus. The two *ltm* genes, *ltmE* and *ltmJ*, are shown as green arrows, the exons of the *ltm* genes are blue boxes under the gene. The lambda clone, λCY324, is shown as an arrow. The primers, lol205 and lol206, used for amplification of the probe fragment are shown above the gene. The fragment used as a probe to isolate the lambda clones and for the Southern analysis in Fig. 3.40 is shown as a red box above the restriction enzyme map. The hybridisation with fragments *ltmE* and a fragment spanning *ltmE-ltmJ* was used to extend the map towards the left by IPCR using the restriction enzymes *Cla*I, *Xba*I and *Hin*dIII.

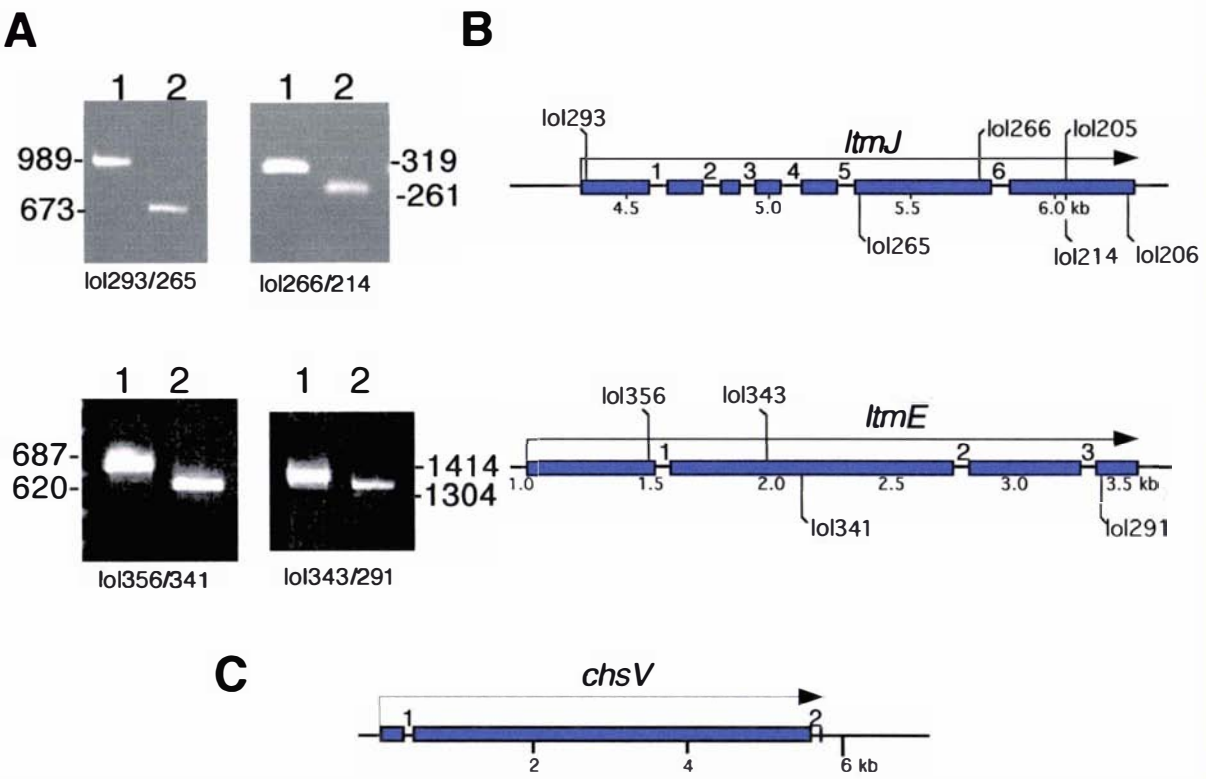


Figure 3.45 Gene structures of the *ltm* cluster 3 genes

Gene structures of the *ltm* cluster 3 genes. (**A**) RT-PCR confirming placement of introns. Lane (1) Lp19 genomic DNA; (2) cDNA from Lp19-infected Nui perennial ryegrass. The RT-PCR products were purified and sequenced with gene specific primers. Primer combinations used are indicated under each gel picture. The sizes of the PCR products are indicated in bp. (**B**) Gene structures showing introns and placement of primers used for RT-PCR analysis. The exons are represented by blue boxes and the primers are positioned above or below the sequence for forward or reverse direction respectively. The approximate postion of the gene within the cluster is indicated as kb. (**C**) A schematic diagram of the *chsV* gene.

RT-PCR analysis was performed with cDNA that was reverse transcribed using mRNA from Lp19-infected Nui ryegrass. The PCR reaction conditions were: 5 ng of genomic DNA or 5 μL of diluted cDNA, 1x *Taq* polymerase buffer (Roche), 50 μM each dNTP, 200 nM each primer, 0.5 U *Taq* polymerase in a reaction volume of 25 μL. The PCR amplification conditions were as follows: 1 cycle of 94^{o}C for 2 min; 35 cycles of 94^{o}C for 15 s, 60^{o}C for 30 s, 72^{o}C for 1 min; followed by one cycle of 72^{o}C for 10 min. The extension time was adjusted to 1 min/kb for larger fragments.

endophyte biomass *in planta* is approximately 1% (Sections 3.5; 3.6.2; Tables 3.4, 3.7). Given that expression of *ltmG*, *ltmM* and *ltmK* was highly up-regulated *in planta* (Fig. 3.13), the other genes involved in lolitrem biosynthesis were also expected to follow a similar expression pattern. Random primed cDNA pools were made from mRNA of *N. lolii* Lp19-infected Nui perennial ryegrass and *E. festucae* Fl1- infected meadow fescue, and Lp19 and Fl1 grown in liquid culture. The cDNA pools from the two endophyte growth conditions, *in planta* and in culture, were diluted to levels where the endophyte *tub2* sequences were amplified to similar levels (Fig. 3.46), thereby adjusting the levels of cDNA from endophytes grown in culture to similar levels to those of the endophyte *in planta*. A dilution series of the cDNA synthesised from mRNA of endophyte grown in culture and cDNA synthesised from mRNA of endophyte-infected ryegrass diluted 1/10 was used as templates for the amplification of *tub2* with primers T1.1 and T1.2 (Fig. 3.46). Equivalent expression levels of the *tub2* gene occurred for the cDNA from endophyte-infected ryegrass diluted 1/10 and for a 1/2000 or a 1/400 cDNA dilution from Lp19 or Fl1 culture conditions, respectively (Fig. 3.46C).

The expression pattern of each gene was subsequently compared between the cDNA of endophyte-infected ryegrass and the cDNA from the endophyte alone (Fig. 3.47). The expression of the 10 *ltm* genes have similar transcript levels which indicated that these genes are highly up-regulated *in planta*. No transcript was detected for any of the 10-*ltm* genes from cDNA of endophytes grown in culture. An additional sample, 10-fold more concentrated, of cDNA synthesised from the endophyte culture was included in the comparison to unequivocally show that the expression patterns from the culture condition did not contain *ltm* transcripts. The expression of *chsV* is similar to that of *tub2* where the gene appears to be constitutively expressed *in planta* and in culture. No evidence of *pks* expression is seen in either endophyte-infected plant material or in culture (Fig. 3.47).

The *ltm* genes were all up-regulated *in planta* and were therefore likely to share common regulatory elements. The 5' untranslated regions of the Lp19 *ggs1* gene and each *ltm* gene were analysed for motifs that were over-represented within the *ltm* sequences. Initially, 1001 bases of promoter sequence, including the ATG of each gene, was subjected to BLASTN analysis to determine if there were regions that shared identity between the promoters. The BLASTN analysis revealed three independent blocks of sequence with significant identity between three pairs of *ltm* promoter

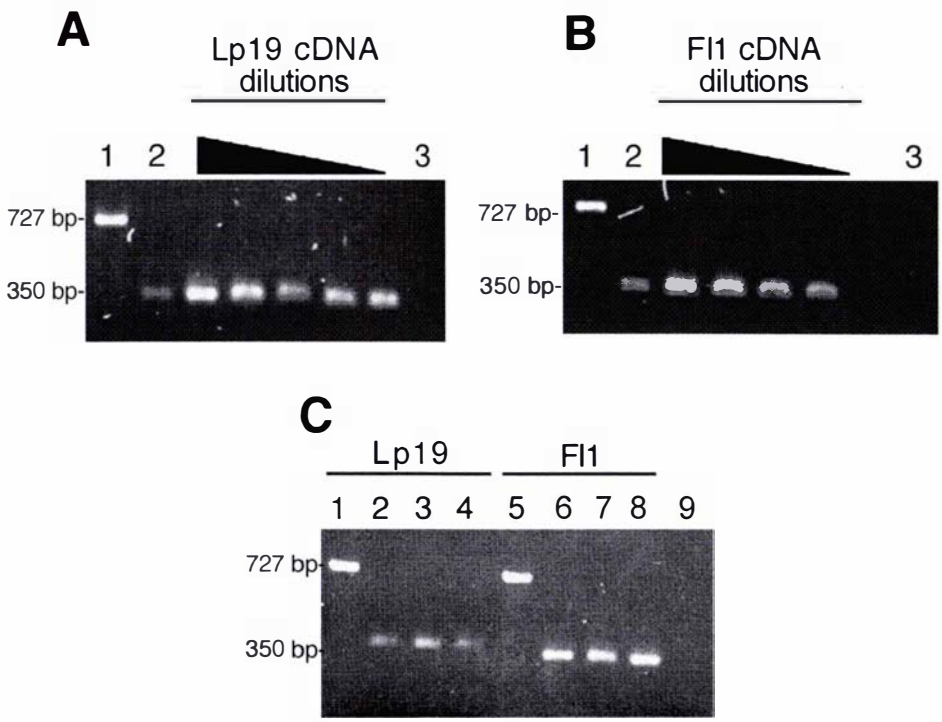


Figure 3.46 PCR analysis to determine cDNA dilutions with equivalent amplification of *tub2*

Amplification of endophyte *tub2* with primers T1.1 and T1.2, to determine cDNA dilutions with equivalent amplification of *tub2*.

(**A**) Amplification of cDNA from Lp19-infected Nui ryegrass with a dilution series of Lp19 cDNA from liquid culture. Lane (1) Lp19 genomic DNA; (2) cDNA of Lp19-infected Nui ryegrass diluted 1/10; (3) water control. The Lp19 cDNA dilutions covered the range 1/100, 1/200, 1/300, 1/400, 1/500.

(**B**) Amplification of cDNA from Fl1-infected meadow fescue ryegrass with a dilution series of Fl1 cDNA from liquid culture. Lane (1) Fl1 genomic DNA; (2) cDNA of Fl1-infected meadow fescue ryegrass diluted 1/10; (3) water control. The Fl1 cDNA dilutions covered the range 1/100, 1/200, 1/300, 1/400, 1/500.

(**C**) 'Fine tuning' the concentration of the cDNA dilutions with respect to the 1/10 cDNA from endophyte-infected ryegrass. Lane (1) Lp19 genomic DNA; (2) cDNA of Lp19-infected Nui ryegrass diluted 1/10; (3) cDNA of Lp19 diluted 1/1000; (4) cDNA of Lp19 diluted 1/2000; (5) Fl1 genomic DNA; (6) cDNA of Fl1-infected meadow fescue diluted 1/10; (7) cDNA of Fl1 diluted 1/400; (8) cDNA of Fl1 diluted 1/500; (9) negative control.

PCR reaction conditions were: 5 ng of genomic DNA or 5 μL of diluted cDNA, 1x *Taq* polymerase buffer (Roche); 50 μM each dNTP; 200 nM each primer, 0.5 U of Taq polymerase in a reaction volume of 25 μL. The PCR amplification conditions were as follows: 94^{o}C for 2 min; followed by 30 cycles of 94^{o}C for 15 s, 60^{o}C for 30 s, 72^{o}C for 1 min; then one cycle of 72^{o}C for 10 min.

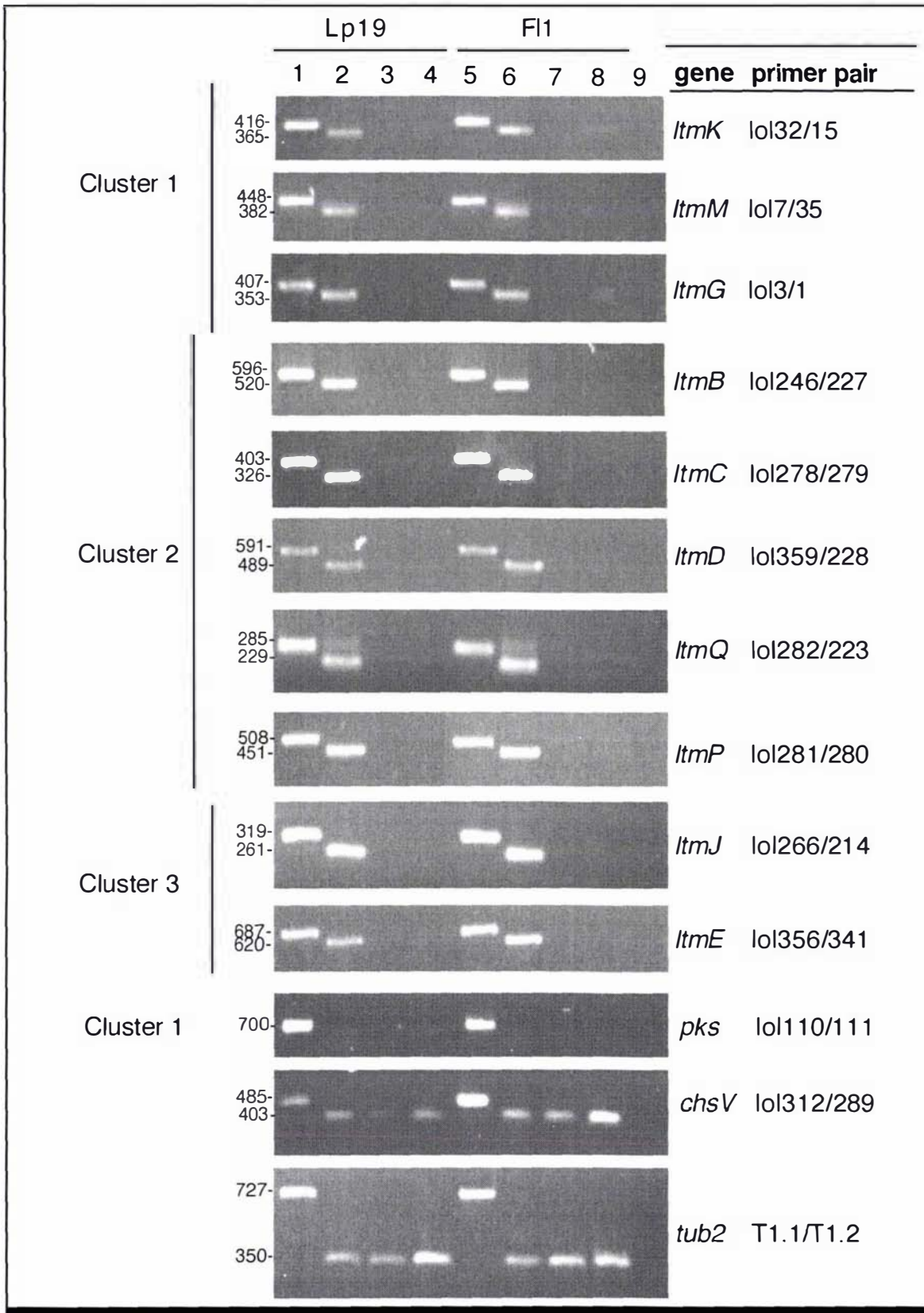


Figure 3.47 Expression analysis of the 10 *ltm* genes

Expression analysis of the 10 endophyte *ltm* genes from clusters 1-3, the pks pseudo gene adjacent to *ltm* cluster 1, chitin synthase (*chsV*) and *tub2* genes. The fragment sizes are in bp. The samples are loaded as follows; Lane (1) Lp19 genomic DNA; (2) cDNA from Lp19-infected Nui perennial ryegrass diluted 1/10; (3) Lp19 cDNA from liquid culture diluted 1/2000 and (4) 1/200; (5) Fl1 genomic DNA; (6) cDNA from Fl1-infected meadow fescue diluted 1/10; (7) Fl1 cDNA from liquid culture diluted 1/400 and (8) 1/40; (9) water negative control.

Each sample was amplifed with the primer pairs indicated, using PCR reaction conditions identical to those described in Fig. 3.12. The PCR amplification conditions were: 94^{o}C for 2 min, followed by 30 cycles of 94^{o}C for 15 s, 60^{o}C for 30 s and 72^{o}C for 1 min; then one cycle of 72^{o}C for 10 min.

sequences. The *ltmG* and *ltmK* promoters shared 89% identity across 38 bases, *ltmM* and *ltmE* shared 81% across 115 bases, and *ltmP* and *ltmD* shared two regions of identity that were 82% and 84% identical across 82 and 139 bases, respectively (Fig. 3.48). No other obvious regions of identity were recognised between the promoter regions. The promoters of each *ltm* gene were further analysed using algorithms from the site of Regulatory Sequence Analysis Tools (Section 2.7) to detect small over-represented regions common to each *ltm* promoter. Initially the *ltm* promoter sequences were subjected to the oligo-analysis, dyad-analysis and consensus algorithms using the available default settings (van Helden, et al., 2000a; van Helden, et al., 1998; van Helden, et al., 2000b). Motifs that were over-represented and therefore had a positive significance index were used as an input into the pattern finding programme dna-pattern, to locate the motifs within the promoter regions of the *ggs1* and *ltm* genes. The final motif screen was limited to those sequences that were present in the majority of the *ltm* genes and were not identified within the *ggs1* promoter (Table 3.13; Fig. 3.48). The putative motifs and the number of times they were identified within the *ltm* promoters are shown in Table 3.13. The dyad CTAN{11}CTA that was detected with the dyad-analysis programme was identified in all the *ltm* promoters apart from *ltmC*. However, *ltmC* did contain a putative element when the search was extended to CTAN{10-12}CTA. The CTA dyads were not detected in *ggs1*.

DNA sequence motifs that had been recognised in the fungal promoters of genes up-regulated during infection or *in planta* from *Ustilago maydis, mig2* (Farfsing, et al., 2005), *Magnaporthe grisea ACE1* (Bohnert, et al., 2004) and *Cladosporum fulvum Avr9* (Snoeijers, et al., 2003) were found in both the *ggs1* and *ltm* promoters (Table 3.13). Therefore, none of these motifs were exclusively identified within the *ltm* promoters.

3.9.4 Functional analysis of *ltmC*

Functional characterisation of *ltmC* was determined by complementation of the *P. paxilli paxC* deletion mutant, ABC83. The ABC83 mutant is blocked early in the paxilline biosynthesis pathway and therefore unable to synthesis indole-diterpenoids (Bryant, Astin and Scott, unpublished). To express *ltmC* in the *P. paxilli* background, the gene was put under the control of the *paxM* promoter in pPN1851 (Section 3.6.5). The sequences of the Lp19 and Fl1 *ltmC* genes are identical; therefore the *ltmC* gene was amplified from Lp19 genomic DNA using the high-fidelity proofreading enzyme

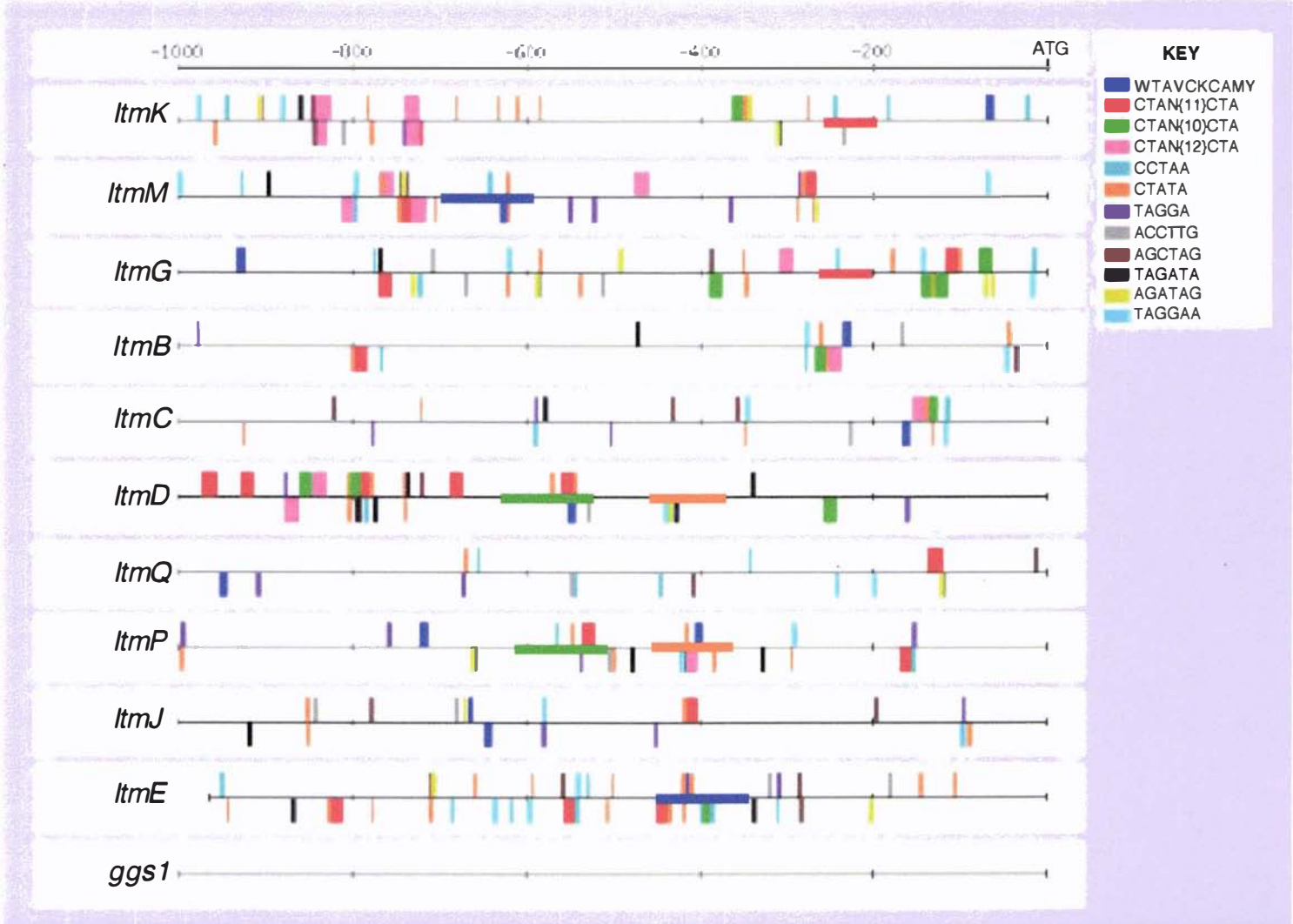


Figure 3.48 Putative regulatory motifs of the *ltm* genes

A schematic map of the putative regulatory motifs identified from the Consensus, Oligo-analysis and Dyad-analysis programmes at the http://rsat.ulb.ac.be/rsat/ site. The diagram was drawn using the programme dna-patterns and feature map. The 5' non-coding regions of *ggs1* and each *ltm* gene was used for the analysis. Motifs are represented as vertical coloured bars either above (top strand) or below (bottom strand) the line. The horizontal coloured bars show the regions of >80% identity between the promoter sequences. The number of the motifs present in each promoter region is shown in Table 3.13.

Table 3.13 Over-represented motifs from the *ltm* genes

DNA Pattern[1]	Programme[2] Reference	Not present[3]	Number of matches identifed in each 5' noncoding region										
			ltmK	*ltmM*	*ltmG*	*ltmB*	*ltmC*	*ltmD*	*ltmQ*	*ltmP*	*ltmJ*	*ltmE*	ggs1
WTAVCKCAMY	consensus	*ggs1*	1	1	1	1	1	1	1	2	2	1	0
$CTAN_{11}CTA$	dyad	*ggs1, ltmC*	1	2	2	1	0	5	2	2	1	3	0
$CTAN_{10}CTA$		*ggs1, ltmM, ltmP*	1	0	4	1	1	3	4	0	0	1	0
$CTAN_{12}CTA$		*ggs1, ltmJ, ltmE*	4	4	1	1	1	2	1	1	0	0	0
CCTAA	Oligo-5	*ggs1*	4	3	4	2	2	1	4	2	2	8	0
CTATA	Oligo-5	*ggs1*	9	7	8	4	5	7	8	6	4	15	0
TAGGA	Oligo-5	*ggs1*	4	8	3	3	5	3	3	6	3	4	0
ACCTTG	Oligo-6	*ggs1, ltmM*	2	0	3	1	1	1	3	1	2	2	0
AGCTAG	Oligo-6	*ggs1*	2	1	1	1	3	2	1	1	2	3	0
TAGATA[4]	Oligo-6	*ggs1*	2	2	2	1	1	5	1	3	1	3	0
AGATAG	Oligo-6	*ggs1, ltmB, ltmC*	3	2	5	0	0	1	5	1	1	2	0
TAGGAA	Oligo-6	*ggs1, ltmJ*	2	3	3	2	2	1	3	2	0	3	0
CAA	Farfsing et al. 2005		5	20	14	31	28	12	22	20	24	8	40
CCAMM	Farfsing et al. 2005		1	5	8	5	6	3	4	3	2	2	12
CCAMC	Farfsing et al. 2005	*LtmK, ltmD, ltmJ*	0	1	2	1	3	0	2	2	0	1	5
MNMNWNCCAMM	Farfsing et al. 2005	*ltmK, ltmB, ltmJ, ltmE*	0	1	1	0	1	1	1	1	0	0	3
TAGATA[4]	Snoeijers et al. 2003	*ggs1, ltmK,M,G,B,C,D,Q,P,J,E*	0	0	0	0	0	0	0	0	0	0	0
GAL4 [5]	Bonhert et al. 2004	*ltmK, G, B, D, Q, P*	0	1	0	0	1	0	0	0	5	1	0

[1]The ambiguity code is as follows, W=A or T; V=A, C or G; K=T or G; M=A or C; Y=C or T

[2]The programme from http://rsat.ulb.ac.be/rsat/ used to identify the sequence. The programmes Oligo-5 and Oligo-6 selected 5 and 6 nucleotide patterns respectively.

[3]The genes whose 5' untranslated regions did not contain the motif.

[4]The TAGATA motif was identified by oligo-analysis. The motif identified in Snoeijers et al. 2003 is an overlaping inverted repeat of TAGATA which was not contained in any promoter elements.

[5]The screen for potential Gal4 domains used the motif CGGN{4,12}CCG.

Platinum *Pfx* (Invitrogen) with primers lol235 and lol236. These primers incorporate *Nco*I and *Eco*RI restriction enzyme recognition sites, respectively. The 1242-bp PCR fragment, containing the *ltmC* gene and 109-bp of 3' untranslated region, was digested with *Nco*I and *Eco*RI and directionally cloned into pPN1851, resulting in plasmid pCY34 (Fig. 3.49). The *ltmC* gene was fused to the *paxM* promoter at the ATG translational start site using the restriction endonuclease *Nco*I. The translational fusion that results in creating an *Nco*I site in the *ltmC* gene caused a single base change where the second codon of *ltmC* has a conservative replacement of threonine in the wild-type gene to alanine in the fused gene. A 3.5 kb *Hin*dIII fragment from λCY315 was cloned into a pUC118 vector resulting in plasmid pCY66 (Fig. 3.49). This 3.5 kb *Hin*dIII fragment contained the complete Lp19 *ltmC* gene under the control of its native promoter. Protoplasts of ABC83 were transformed with pII99 and pJA8, containing an endogenous *paxC* fragment, or co-transformed with pCY34 and pII99, or pCY66 and pII99, and transformants selected on geneticin. Approximately 5-10 stable *P. paxilli* ABC83 transformants were colony-purified and subsequently screened by TLC analysis for their ability to produce paxilline (Fig. 3.50).

TLC analysis of the wild-type *P. paxilli* indole-diterpenoid extraction showed intense green bands that have the same Rf as paxilline, paspaline and 13-desoxypaxilline (Fig. 3.50). The ABC83 *paxC* mutant, used for the transformations, was unable to produce any indole-diterpene (Fig. 3.50). The ABC83 transformants containing pII99 are unable to complement the *paxC* mutation and are therefore paxilline negative (Fig. 3.50; samples ABC283-1 and 2). The ABC83 transformants co-transformed with plasmids pII99 and pCY66 with the Lp19 *ltmC* gene under the control of the native Lp19 promoter are unable to complement the *paxC* mutation and are paxilline negative (Fig. 3.49; samples ABC383-1 - 5). All five ABC83 transformants containing the endogenous *paxC* gene on plasmid pJA8 were able to complement the *paxC* deletion phenotype (Fig. 3.50; samples ABC483-1 - 5). Seven of the 10 transformants containing *ltmC* under the control of the *paxM* promoter are able to produce paxilline (Fig. 3.50; samples ABC583-1 - 10). The TLC analysis was confirmed by HPLC analysis (Appendix 5.4). These data confirmed that *ltmC* is a functional orthologue of *paxC*.

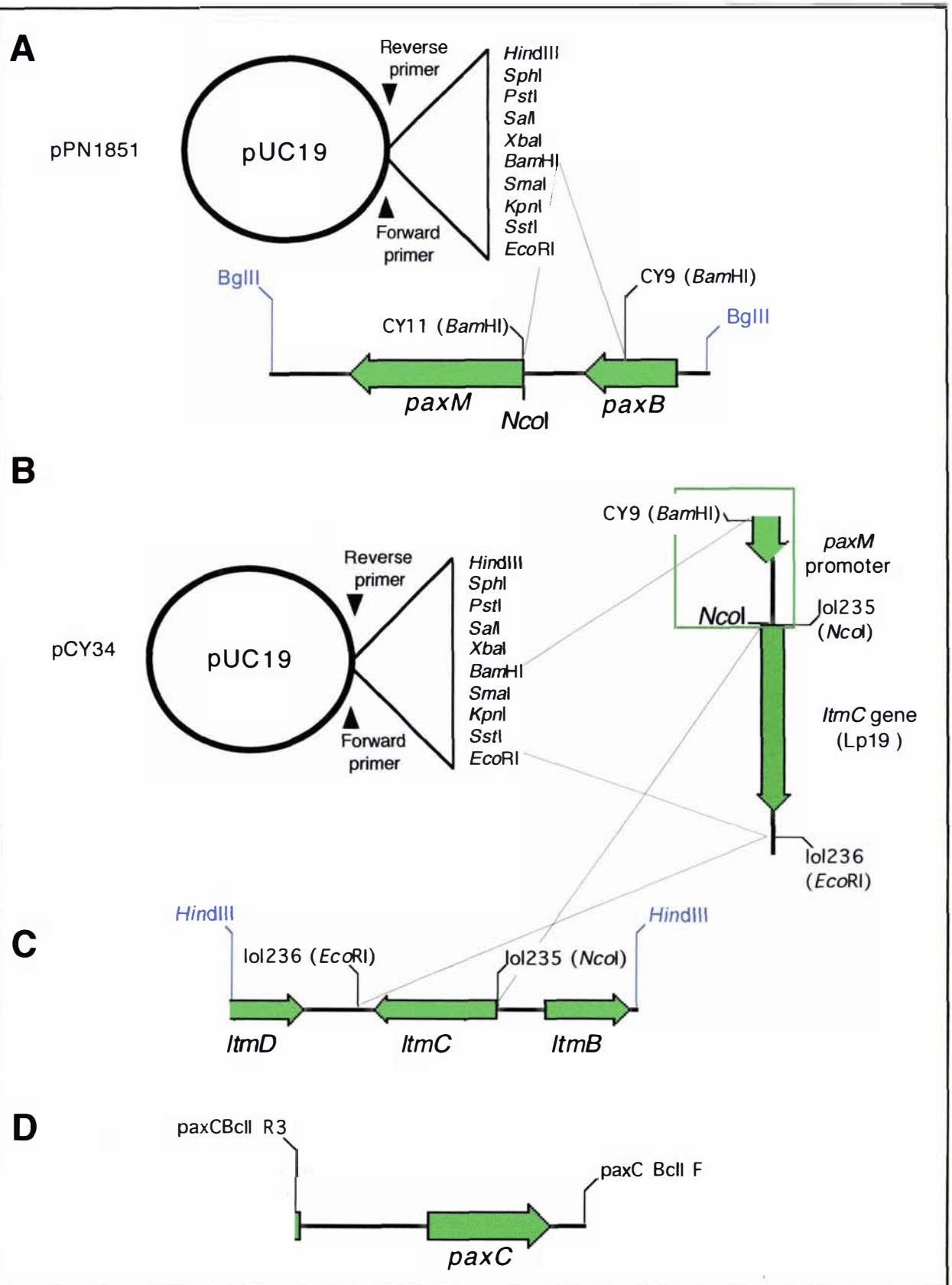


Figure 3.49 Constructs for complementation of the *paxC* deletion mutant

Making the constructs for complementation of the *paxC* deletion mutant. (**A**) The pPN1851 construct as shown in Fig. 3.29. (**B**) The pCY34 construct. The *ltmC* gene from Lp19 was amplified with primers lol235 and lol236, digested with *Nco*I and *Eco*RI and subsequently cloned into pPN1851. The *paxM* promoter is in a green box. (**C**) The 3.5 kb *Hin*dIII fragment from Lp19 containing *ltmC* was cloned into pUC118 resulting in pCY66. The pCY66 plasmid was used with pII99 in a co-transformation of ABC83 protoplasts. (**D**) The 2.5 kb *Bcl*I fragment from *P. paxilli* containing *paxC* was cloned into pII99 resulting in pJA8 (Astin, unpublished).

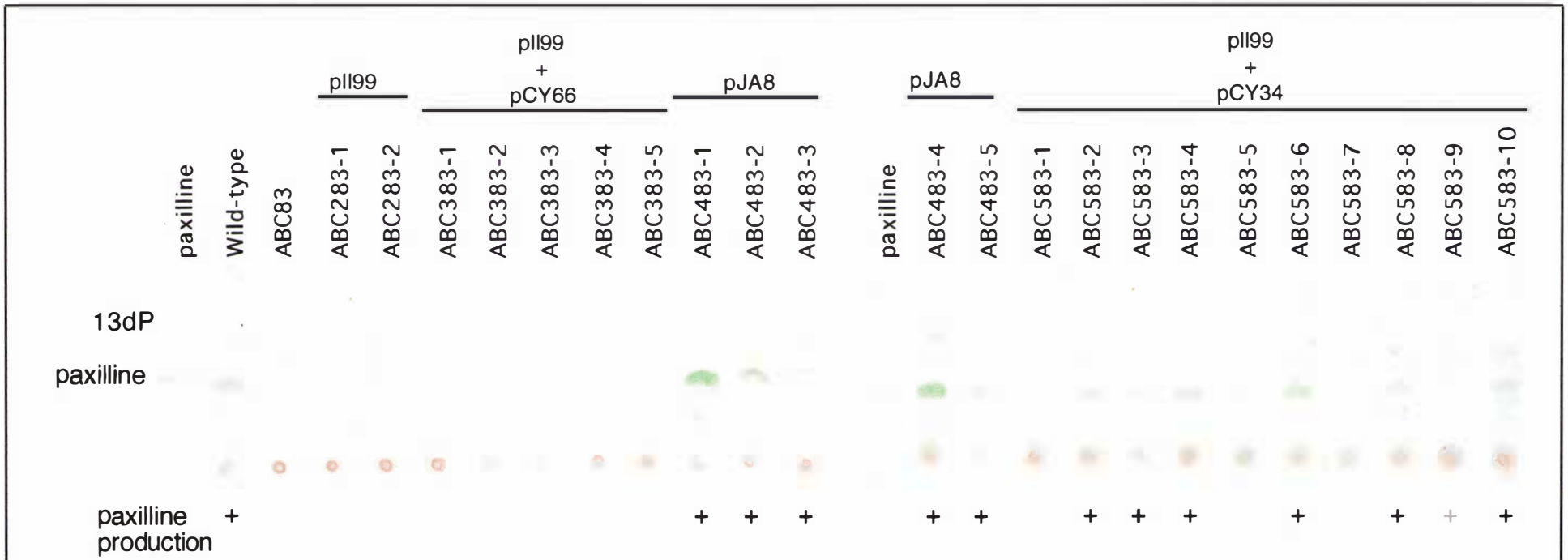


Figure 3.50 TLC analysis of *paxC* complementations

TLC analysis of *paxC* complementation transformants. Indole-diterpenes were extracted from mycelium (Section 2.11) grown for 7 days in CDYE + trace elements (Section 2.2.2, 2.2.7) at 28°C. All plasmids were used to transform the *paxC* deletion mutant, ABC83. The plasmids were as follows; pII99; pCY66 withLp19 *ltmC* under the control of its native promoter; pJA8 with *paxC* under the control of its native promoter; pCY34 with the Lp19 *ltmC* gene under the control of the *paxM* promoter. The + under the TLC plate indicates the presence of a green band identical in Rf to the paxilline standard, whereas the + indicates possible paxilline production. 13dP = the mobility of paspaline and 13-desoxypaxilline.

3.9.5 *ltm* clusters 1, 2 and 3 form a large platform

Sequence analyses of the *ltm* cluster 2 lambda clones λCY301, λCY305 and λCY346 showed linkage to lambda clone λCY324 containing *ltm* cluster 3 (Section 3.9.2; Fig. 3.51). The two clusters are separated by ~16 kb of repetitive AT-rich (69% AT) DNA that contains a series of direct repeats as shown in Fig 3.51. Definition of the repeating units between *ltm* clusters 2 and 3 resolved the hybridisation patterns found within the λCY301 fragment (Fig. 3.43). The expected hybridisation patterns to the 301 probe (Section 3.9.1) would be to 18.3 kb *Eco*RI, 10.7 kb *Hin*dIII and 18.4 kb *Sst*I fragments, with each hybridising fragment containing two binding sites of the probe (Fig. 3.43; 3.51). The actual hybridisation pattern (Fig. 3.43) had additional hybridising fragments, which suggested that the repeat is present elsewhere in the genome. The nucleotide sequence of the gap between *ltm* clusters 2 and 3 was also compared to sequences of λCY224, λCY226, λCY227 and λCY229 that were identified from the library screen with the Rua probe (Section 3.7.2). A 0.8-kb sequence was common to λCY229, λCY227 and the region between *ltm* clusters 2 and 3 (Fig. 3.51). However, lambda clone λCY227 contained only 339 bp of this region.

To determine whether cluster 1 was contained on the same chromosome as clusters 2 and 3, the fragments *ltmM*, *ltmB*, and *ltmJ* from *ltm* clusters 1, 2 and 3, respectively, were hybridised to endophyte chromosome separations. The three genes are contained on the same 2.5-Mb and 6.0-Mb chromosomes from Lp19 and Fl1, respectively (Fig. 3.52). The *N. lolii* strain Lp7 contains the three genes on a 4.6-Mb chromosome. The *ltm* clusters 1 and 2 are contained on 2.5 Mb chromosome in *N. lolii* strain Lp5, but the hybridisation with *ltmJ* was to a larger chromosome of ~3.5 Mb (Fig. 3.52). The *N. lolii* strain, AR1, only hybridises to *ltmM* and *ltmB* from clusters 1 and 2. The Lp1 strain from *Neotyphodium* LpTG-2 has not hybridised with the three probes, due to the quality of the DNA. As expected from previous data (Fig. 3.4, 3.40) E8 did not hybridise with the *ltm* genes (Fig. 3.52). Fr1, an *E. festucae,* did not hybridise with the *ltmM* or *ltmB* probes and had faint hybridisation with *ltmJ* to a 4.5-Mb chromosome. However, the *ltmJ* probe contained the highly conserved haem-binding domain, which could cross hybridise with other cytochrome P450 monooxygenases.

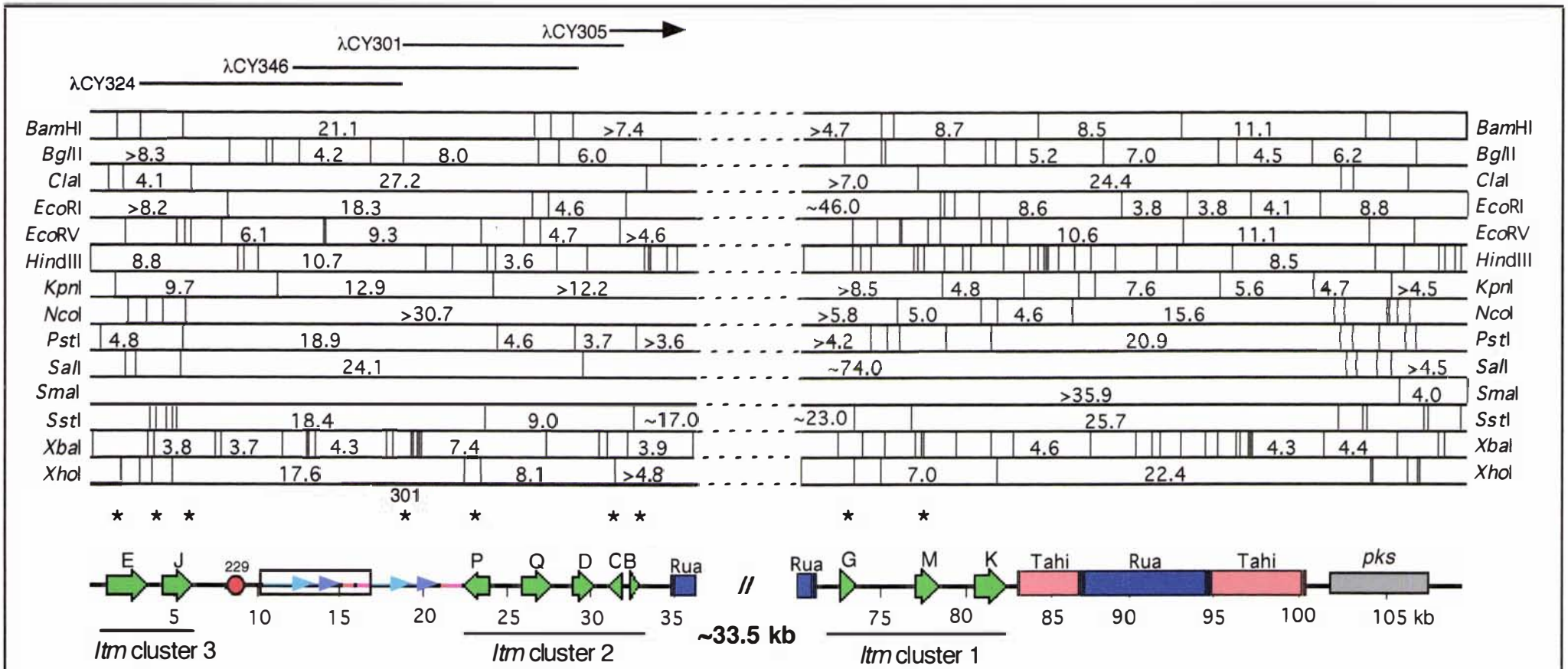


Figure 3.51 An *Itm* gene platform

Linking the *Itm* gene clusters 1, 2 and 3 as a large platform. The genes are represented by green arrows, the retrotransposons are represented by red and blue boxes for Tahi and Rua respectively. The approximate distance across the platform is represented as kb. The • are regions that have been used for hybridisation analysis. The lambda clones are black lines above the restriction enzyme map. The uncloned region between cluster 1 and clusters 2 and 3 is estimated from Figure 3.53 based on the *Eco*RI, *Sal*I and *Sst*I hybridisation patterns. The repeat units between clusters 2 and 3 are 3 kb, 1.8 kb, 0.44 kb and 0.56 kb. A single repeat is surrounded by a black box. The sequence common to λCY229 and λCY227 is shown as a red circle.

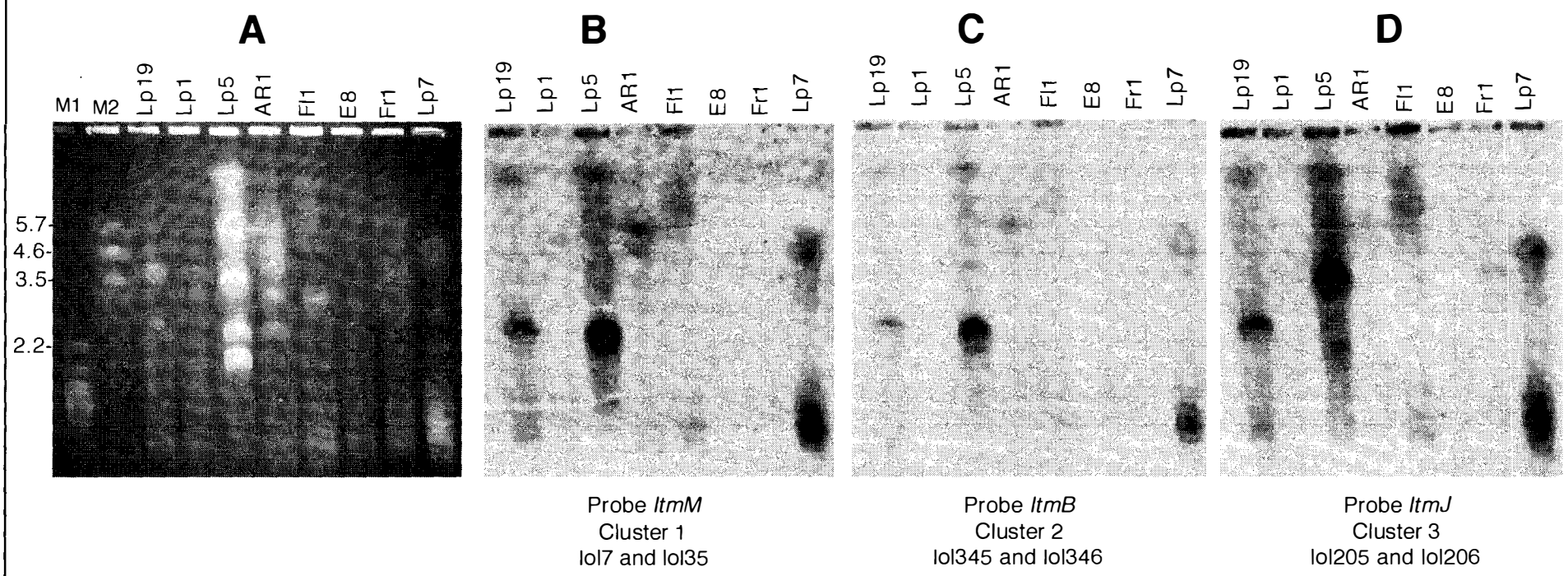


Figure 3.52 The chromosomal location of the *ltm* gene clusters 1, 2 and 3

The chromosomal locations of *ltm* genes from clusters 1, 2 and 3. (**A**) Chromosomal separation of *N. lolii* strains Lp19, Lp5, AR1 and Lp7, *E. festucae* strains Fl1 and Fr1, *E. typhina* strain E8 and *Neotyphodium* LpTG-2 strain Lp1. Lane M1 = *S. cerevisiae* chromosomal standard (BioRad), M2 = *S. pombe* chromosomal standard (BioRad). The size standards are indicated in Mb. The electrophoresis conditions are the same as in Fig. 3.6. Autoradiograph of the gel in **A** hybridised with ^{32}P-labelled (**B**) *ltmM* probe; (**C**) *ltmB* probe; (**D**) *ltmJ* probe. The primers used for probe amplification and the cluster location of each fragment is indicated under each blot.

To determine if *ltm* cluster 1 was linked to *ltm* clusters 2 and 3, Southern analysis of digests with the rare base cutter, *Not*I, and the restriction enzymes *Eco*RI, *Sal*I and *Sst*I were hybridised with *ltmG*, *ltmB* and *ltmJ* probes (Fig. 3.53). The *ltmG* and *ltmB* fragments both hybridised to ~48 kb *Eco*RI fragments from Lp19, Fl1 and Lp1. The hybridisation with *ltmG* and *ltmB* to *Sal*I digests shows that these two fragments are contained on a ~74 kb *Sal*I fragment in Lp19 and Lp1 and on a 50 kb fragment in Fl1. The difference in size of the *Sal*I fragments between Lp19 and Fl1 is due to the absence in Fl1 of the retrotransposon platform between *ltm* cluster 1 and the *pks* (Section 3.7.3). The *ltmG* and *ltmB* probes hybridised to two different-sized *SstI* fragments of 23 kb and 17 kb in all three strains (Fig. 3.53). Based on the sizes of the *ltmG* and *ltmB* hybridising fragments, cluster 1 is separated from cluster 2 by ~35 kb (Fig. 3.51). The Lp1 strain was used as a control to confirm that Lp1 and Lp19 have similar-sized hybridisation bands as those seen in Fig. 3.36.

The *E. festucae* Fl1 *Not*I digests contained a hybridising band of ~120 kb with the probes *ltmJ*, *ltmB* and *ltmG*, suggesting *ltm* clusters 1, 2 and 3 are contained on the same fragment (Fig. 3.53). The *N.* spp. Lp1 strain hybridised to a band of ~120 kb with the *ltmG* and *ltmB* probes but no hybridising band was detected with the *ltmJ* probe (Fig. 3.53). The intensities of the *N.* spp. Lp1 hybridising bands are lower, when compared to *N. lolii* Lp19 and *E. festucae* Fl1, as Lp1 is a hybrid (*E. festucae* x *E. typhina* complex) and therefore has a genome twice the size of these strains. Lp19 *Not*I digests contained two hybridising bands of ~170-180 kb to each of the three probes (Fig. 3.53E). This suggested that not only are the three clusters linked in Lp19 but this region is also duplicated (Fig. 3.53). The hybridisation data confirms that *ltm* clusters 1, 2 and 3 are linked, and therefore, the *ltm* genes are contained together as a platform in Lp19 and Fl1. Following this result, a re-analysis of earlier data (Fig 3.4, 3.14, 3.32, 3.33, 3.39) was carried out which showed Lp19 hybridisations had not detected a duplication of the *ltm* region. However, *ltm* genes hybridised to two Lp19 chromosomes (Fig. 3.35; 3.52), the obvious 2.5 Mb chromosome and a much fainter hybridisation to that of the largest chromosome at ~10 Mb. A similar hybridisation pattern is seen with the Lp5 chromosomes (Fig. 3.52).

Probes to six regions of the *ltm* platform were hybridised using standard hybridization conditions (Section 2.4.7.1) to 14 different restriction enzyme digests of Lp19 genomic DNA (Fig. 3.54). Of the 14 digests screened, only that of *Pst*I show a different banding

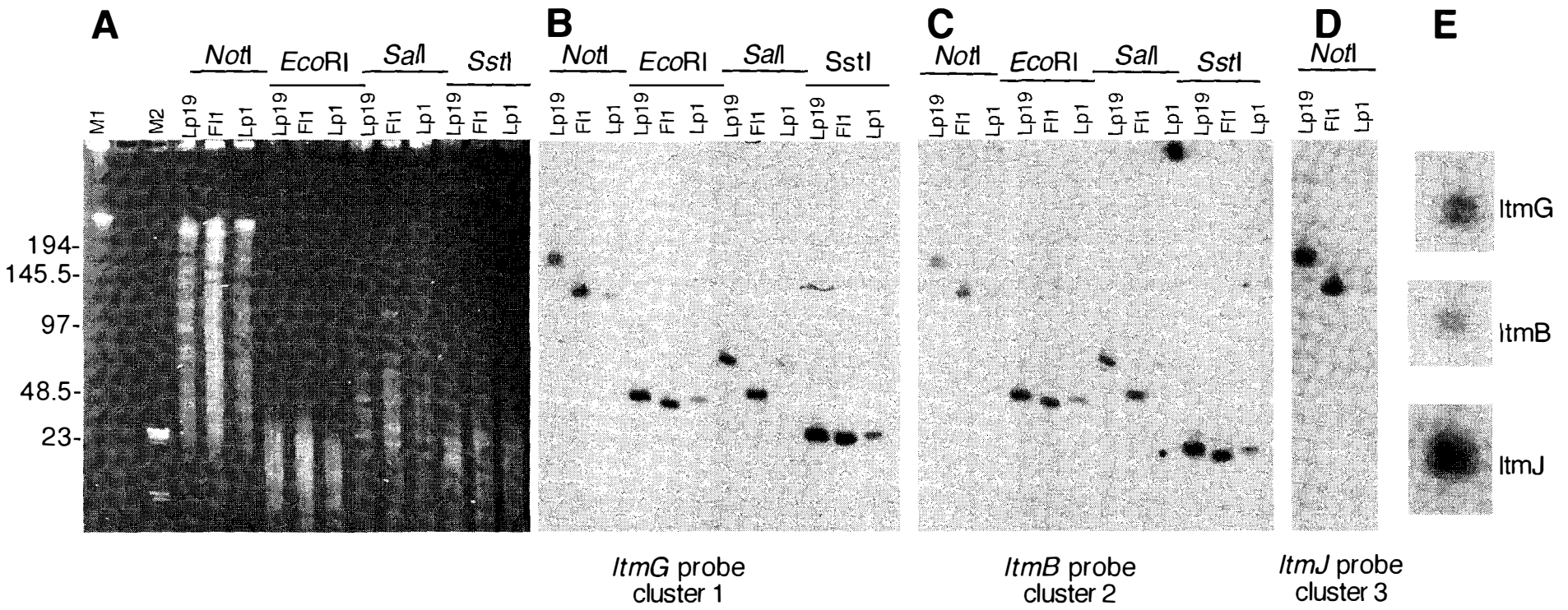


Figure 3.53 Linkage of *Itm* cluster 1 and 2 established by Southern analysis

Southern analysis to show linkage of *Itm* clusters 1, 2 and 3. (**A**) An ethidium bromide stained gel of genomic digests. (**B-E**) Autoradiographs of hybridisations of **A** with ^{32}P-labelled fragments of *ItmG*, *ItmB* and *ItmJ*, from clusters 1, 2 and 3 respectively, to *Not*I, *Eco*RI, *Sal*I and *Sst*I restriction enzyme digests of Lp19, Fl1 and Lp1. Lp1 was used as a control to confirm that Lp19 and Lp1 have similar-sized hybridisation bands as those seen in Fig. 3.36. Lane M1, Lambda ladder (BioRad); lane M2 lambda *Hin*dIII DNA standard (Invitrogen). Size standards are shown in kb. Hybridisations are with (**B**) *ItmG* (amplified with primers lol79 and lol1), (**C**) *ItmB* (amplified with primers lol345 and lol346), (**D**) *ItmJ* (amplified with primers lol205 and lol206). (**E**) Enlargements of the Lp19 *Not*I lanes hybridised with each probe to show the hybridising band was a doublet.

Conditions for DNA separation: 1.0 % (w/v) Molecular biology grade agarose in 0.5x TBE at 200 volts, with the following ramped switch times 2-10 s for 18.4 h, 10-16 s for 4.6 h, at 14^{O}C.

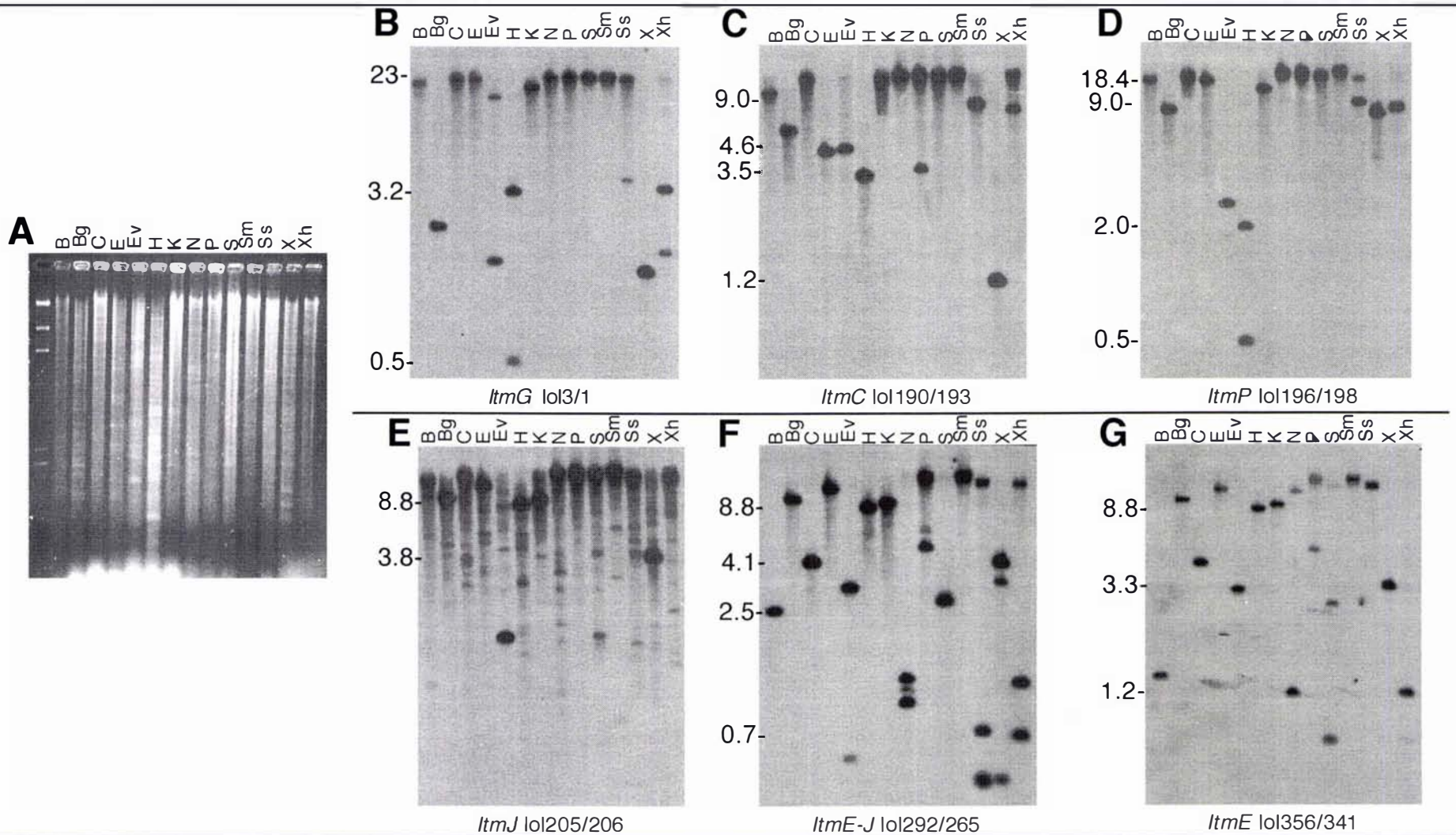


Figure 3.54 Southern analysis of the Lp19 *ltm* gene platform

Southern analysis of the Lp19 *ltm* gene platform. (**A**) The ethidium bromide stained gel of digested Lp19 genomic DNA. Autoradiographs of **A** hybridised with ^{32}P-labelled probes of **(B)** *ltmG* from *ltm* cluster 1; (**C**) *ltmC* and (**D**) *ltmP* from *ltm* cluster 2; (**E**) *ltmJ*, (**F**) *ltmE-J* and (**G**) *ltmE* from *ltm* cluster 3. The hybridisation of *ltmJ* has background due to the probe being located across the highly conserved haem-binding domain and that this was the first probe hybridised to this blot. The restriction enzymes used were: B=*Bam*HI, Bg=*Bgl*II, C=*Cla*I, E=*Eco*RI, Ev=*Eco*RV, H=*Hin*dIII, K=*Kpn*I, N=*Nco*I, P=*Pst*I, S=*Sal*I, Sm=*Sma*I, Ss=*Sst*I, X=*Xba*I, Xh=*Xho*I. The *PstI* digest may be incomplete. The location of the probes can be found in Fig. 3.11, 3.41, 3.44 and 3.50. The fragments sizes are in kb.

pattern to what was expected, based on the Lp19 genomic map. However, it is more probable that this sample was not digested to completion. The high background found with the hybridisation with the *ltmJ* fragment results from the presence of the highly conserved haem-binding domain of cytochrome P450 monooxygenases within this probe. The background hybridisation shown in Figure 3.54E is more dramatic than was seen with hybridisations of the *ltmJ* probe to other filters (Fig. 3.53; 3.55; 3.56) because the *ltmJ* fragment was the first probe used with this filter (Fig. 3.54E).

3.9.6 The *ltm* platform is genetically unstable

Previous Southern hybridisation data (Fig. 3.52) indicated that within the *E. festucae* and *N. lolii* group there has been rearrangement of the *ltm* gene platform. This is especially evident in *N. lolii* Lp5 where clusters 1 and 2 are on a different chromosome to cluster 3; that *N.* spp. Lp1 and *N. lolii* AR1 are missing *ltm* cluster 3; and that the *E. festucae* Fr1 strain, which is closely related to Fl1 is missing all three clusters (Fig. 3.52). To assess this further, the genomic arrangement of the *ltm* platform was screened in closely related strains with known indole-diterpene phenotypes. The lolitrem producers were *N. lolii* Lp19 and Lp5, and *E. festucae* Fl1. *N. lolii* Lp14 produces janthitrems, compounds structurally related to the lolitrems (Tapper and Lane, unpublished). The lolitrem non-producers were *N. lolii* AR1, *Neotyphodium* spp. Lp1, and *E. typhina* E8. The chemotype of *E. festucae* E189 is unknown. These isolates were screened for the presence of genes *ltmP*, *ltmJ* and *ltmE* as strains had previously been shown by Southern analysis to have differences in this region. As very little nucleotide sequence diversity is found between the asexual *N. lolii*, Lp19, and sexual *E. festucae*, Fl1, across the *ltm* genes in cluster 1, (Section 3.7.3), standard hybridisation conditions were used.

The *ltmP* probe hybridised to seven of the eight strains screened (Fig. 3.55). E8 is the only strain negative for *ltmP* hybridisation. The *ltmP* probe contains an *Sst*I site and therefore hybridises to two fragments with a 9-kb hybridising band common to the seven strains that contain *ltmP*. This band in Lp19 contains the genes *ltmC*, *ltmD*, *ltmQ*, and partial sequences of *ltmB* and *ltmP* (Fig. 3.55). Sequence diversity is seen amongst the seven strains that contained *ltmP* as the bands of the *Eco*RI-digested DNA and the second *Sst*I hybridising fragment were of varying sizes. Lp5 has two copies of *ltmP* seen clearly as two hybridising bands in the *Eco*RI-digested DNA.

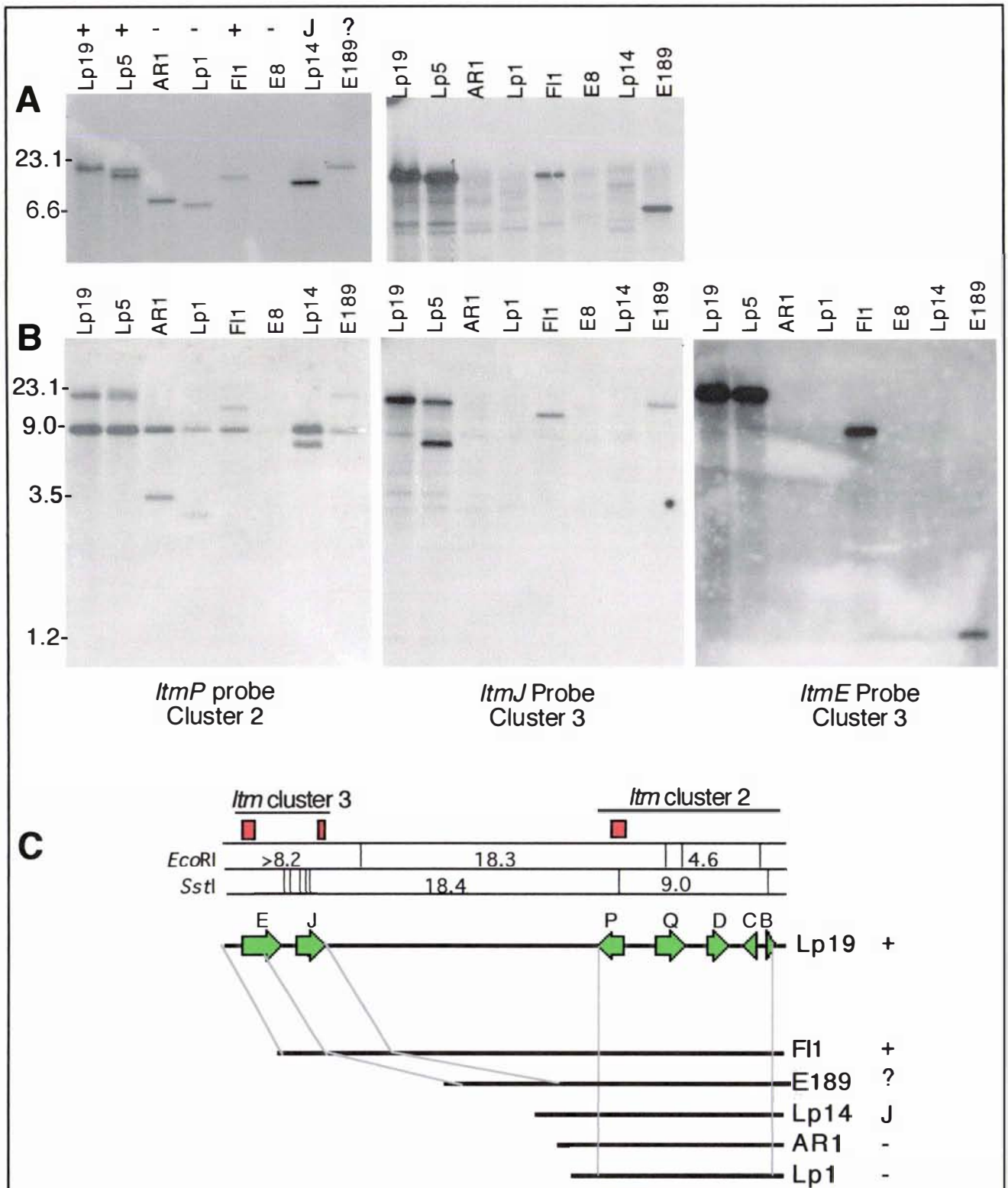


Figure 3.55 Southern analysis of endophyte strains hybridised with genes from *Itm* clusters 2 and 3

Autoradiographs of Southern analysis of (**A**) *Eco*RI-digested; (**B**) *Sst*I-digested, *N. lolii* strains Lp19, Lp5, AR1 and Lp14, *E. festucae* strains Fl1 and E189, *Neotyphodium* species Lp-TG2 strain Lp1 and *E. typhina* strain E8 hybridised with ^{32}P-labelled *ItmP* amplified with primers lol196 and lol198; *ItmJ* amplified with primers lol205 and lol206, and *ItmE* amplified with primers lol356 and lol341. The sizes of the hybridising bands are shown in kb. (**C**) A schematic map of the *Itm* cluster 2 and 3 region showing the approximate deletions in Lp14, Lp1 and AR1 as determined by Southern analysis. The position of the probes are shown as red boxes in **C**. The lolitrem B phenotype of each isolate is indicated in **A** and **C** as follows: + = lolitrem B producer; - = lolitrem B non-producer; J = produces janthitrems an indole diterpene; ? = phenotype unknown.

The *ltmJ* probe hybridises to four (Lp19, Lp5, FII and E189) of the eight strains screened (Fig. 3.55). *E. typhina* E8, *N.* spp. Lp1, *N. lolii* AR1 and *N. lolii* Lp14 are all negative for *ltmJ* hybridisation. *N. lolii* Lp5 contains two copies of *ltmJ*, one that hybridises to an *Sst*I fragment the same size as *N. lolii* Lp19 (~18 kb) and a second 8.5-kb fragment. The *ltmE* probe hybridises to the same four strains, *N. lolii* Lp19, *N. lolii* Lp5, *E. festucae* FII and *E. festucae* E189, as that of the *ltmJ* hybridisation (Fig. 3.55). *E. typhina* E8, *N.* spp. Lp1, *N. lolii* AR1 and *N. lolii* Lp14 are all negative for *ltmE* hybridisation. *N. lolii* Lp19 and Lp5 have the same-sized *ltmE* hybridising band of ~20 kb. *E. festucae* FII and E189 have *ltmE* hybridising bands of 9 kb and 1.2 kb, respectively. The AT-rich region between clusters 2 and 3 is smaller in *E. festucae* FII and E189 than *N. lolii* Lp19 based on the sizes of the hybridising bands with the *ltmP* and *ltmJ* probes (Fig. 3.55). Absence of *ltmJ* and *ltmE* in *N.* spp. Lp1, *N. lolii* AR1, *N. lolii* Lp14 and *E. typhina* E8 correlated with a lolitrem B negative phenotype, suggesting that these two genes are specific for lolitrem biosynthesis. A schematic diagram of the cluster 2 and 3 regions from strains used in the Southern analysis is shown in Figure 3.55. Attempts have been made by IPCR to isolate the regions from Lp1, Lp14 and AR1 that flank the deletions, but to date this has been unsuccessful.

3.10 Phylogenetic distribution of the *ltm* genes

The distribution of the *ltm* genes within *Epichloë*/*Neotyphodium* endophytes was determined to show whether the genes were absent or inactive in non-producers. A selection of *Neotyphodium* and *Epichloë* strains of known and unknown lolitrem phenotypes were analysed for the presence of the *ltm* genes. Initially a screen with *ltmC* was used, as this gene encodes an enzyme that is known to be involved early in the pathway of indole-diterpene biosynthesis. Eighteen endophyte strains from *E. festucae*, *N. lolii*, *Neotyphodium* spp, *E. typhina*, *E. armarillans*, *N. coenophialum*, *N. siegelii*, *E. brachyelytri* and *E. bromicola* were hybridised with *ltmC* using conditions of low stringency and with *ltmM*, *ltmB*, *ltmP* and *ltmJ* using standard hybridisation conditions (Fig. 3.56). Only the *N. lolii* Lp19 and *E. festucae* FII isolates are known to produce lolitrem B. The *N. lolii* isolate AR1, *N.* sp. LpTG-2 isolate Lp1, *N. coenophialum* isolate Tf28 and *E. typhina* isolate E8 are lolitrem B non-producers. The *N. lolii* isolate Lp14 is capable of producing janthitrems, which are structurally similar to lolitrem B. The lolitrem B phenotype of the remaining isolates used in Figure 3.56 are unknown.

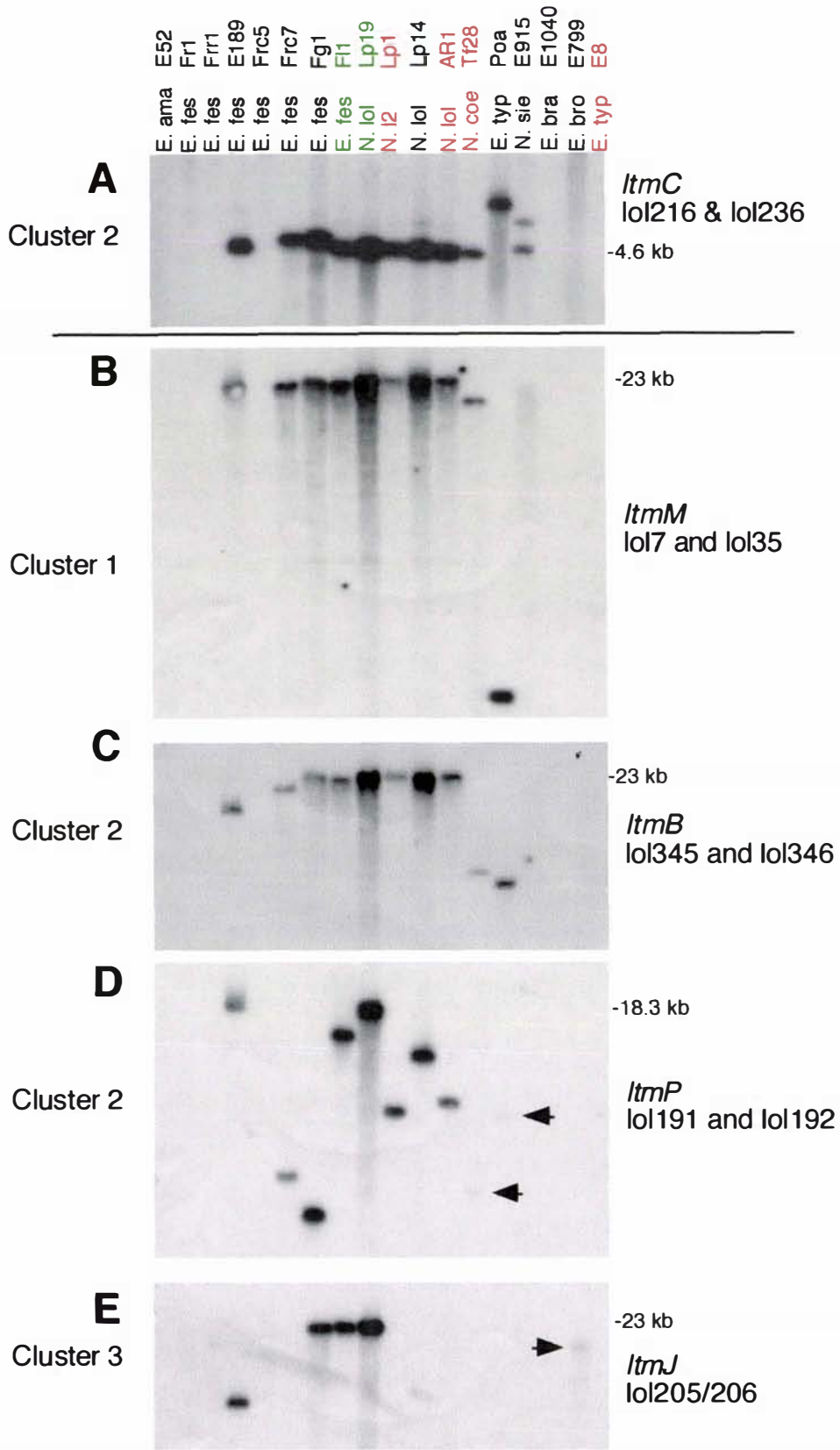


Figure 3.56 Southern analysis of endophyte strains for the *ltm* genes

Southern analysis of *Eco*RI-digested genomic DNA hybridised with ^{32}P-labelled probes. (**A**) Autoradiograph of a low stringency hybridisation (Section 2.4.7) with the complete *ltmC* gene amplified with primers lol216 and 236. A longer exposure failed to detect additional bands. (**B-E**) Autoradiographs from standard hybridisation (Section 2.4.7) with the probes defined adjacent to each blot. The arrows identify bands that were more obvious with a longer exposure. The abbreviations are as follows: E. ama = *E. amarillans*; E. fes = *E. festucae*; N. lol = *N. lolii*; N. l2 = *N.* LpTG-2; N. coe = *N. coenophialum*; E. typ = *E. typhina*; N. sie = *N. siegelii*; E. bra = *E. brachyelytri*; E. bro = *E. bromicola*. The fragment sizes are indicated in kb. The lolitrem B producing isolates are indicated in green text, non-producers in red text and unknown in black text. *N. lolii* Lp14 is able to produce janthitrems an indole diterpene.

The Southern hybridisation data showed that the *E. festucae* isolates formed three groups. E189, Fg1, and Fl1 contained all the *ltm* genes, while Fr1, Frr1 and Frc5 contained none of the *ltm* genes and one strain, Frc7, contained all the *ltm* genes apart from *ltmJ* from cluster 3 (Fig. 3.56). Of the *N. lolii* isolates only Lp19 contained all of the genes, whereas Lp14 and AR1 were missing *ltmJ* from cluster 3 (Fig. 3.56). *N. coenophialum* Tf28, an interspecific hybrid with ancestral origins from *E. festucae*, *E. typhina* complex (ETC) and *Lolium* associated clade (LAE), and *N.* sp. LpTG-2 Lp1 a hybrid with ancestral origins from *E. festucae* and ETC, contain genes from clusters 1 and 2 but are missing *ltmJ* from cluster 3. However, the hybridisation of *ltmP* is very faint in *N. coenophialum* Tf28. *N. coenophialum* Tf28 and *N.* sp. LpTG-2 Lp1 are lolitrem B negative (Christensen, et al., 1993). *E. amarillans* E52, and *E. brachyelytri* E1040 strains were negative for all *ltm* genes (Fig. 3.56). The *E. typhina* isolate E8 showed no hybridisation with the *ltm* genes but Poa, a lolitrem-negative strain, contained genes from cluster 1 and 2 but not cluster 3 (Fig. 3.56). As with *N. coenophialum*, the *E. typhina* Poa hybridisation with *ltmP* was very faint. The quality of the DNA of the *E. bromicola* E799 isolate was poor but there is faint hybridisation with *ltmJ* and *ltmC* (Fig. 3.56). *N. siegelii* E915, a hybrid with ancestral origins from *E. festucae* and *E. bromicola*, only hybridises with the *ltmC* probe and contains two hybridising bands (Fig. 3.56). This hybridisation pattern could be due to partial digestion of the DNA. Of the isolates screened, those that are asexual with an *E. festucae* ancestor have *ltm* genes from clusters 1 and 2 present in their genomes.

The analysis of the *ltm* genes was extended to a larger range of endophytes, covering the 10 sexual *Epichloë* species and a selection of *Neotyphodium* species, using a PCR screen performed with all 10 of the *ltm* genes (Fig. 3.57). Primers used in this screen were designed to regions conserved with known *pax* orthologues or homologues (Appendix 5.1). Two of the genes, *ltmJ* and *ltmE*, were amplified using primers designed for previous PCR and RT-PCR analysis (Appendix 5.1). The geranylgeranyl diphosphate synthase (*ggs1*; Section 3.2) and chitin synthase (*chsV*; Section 3.9.2) genes were used as amplification controls for the DNA as both genes were expected within all endophyte genomes. The PCR with primers to *ggs1* and *chsV* amplified DNA from all the endophyte strains screened (Fig. 3.57). The *E. festucae* isolates, grouped into those with the *ltm* genes, Fg1, Fl1, Frc7 and E189, and those without, Fr1, Frc5, and Frr1, confirming the hybridisation data from Fig. 3.56. Frc7 did not amplify a product with the *ltmJ* or *ltmE* primers while E189 did not amplify with a product with

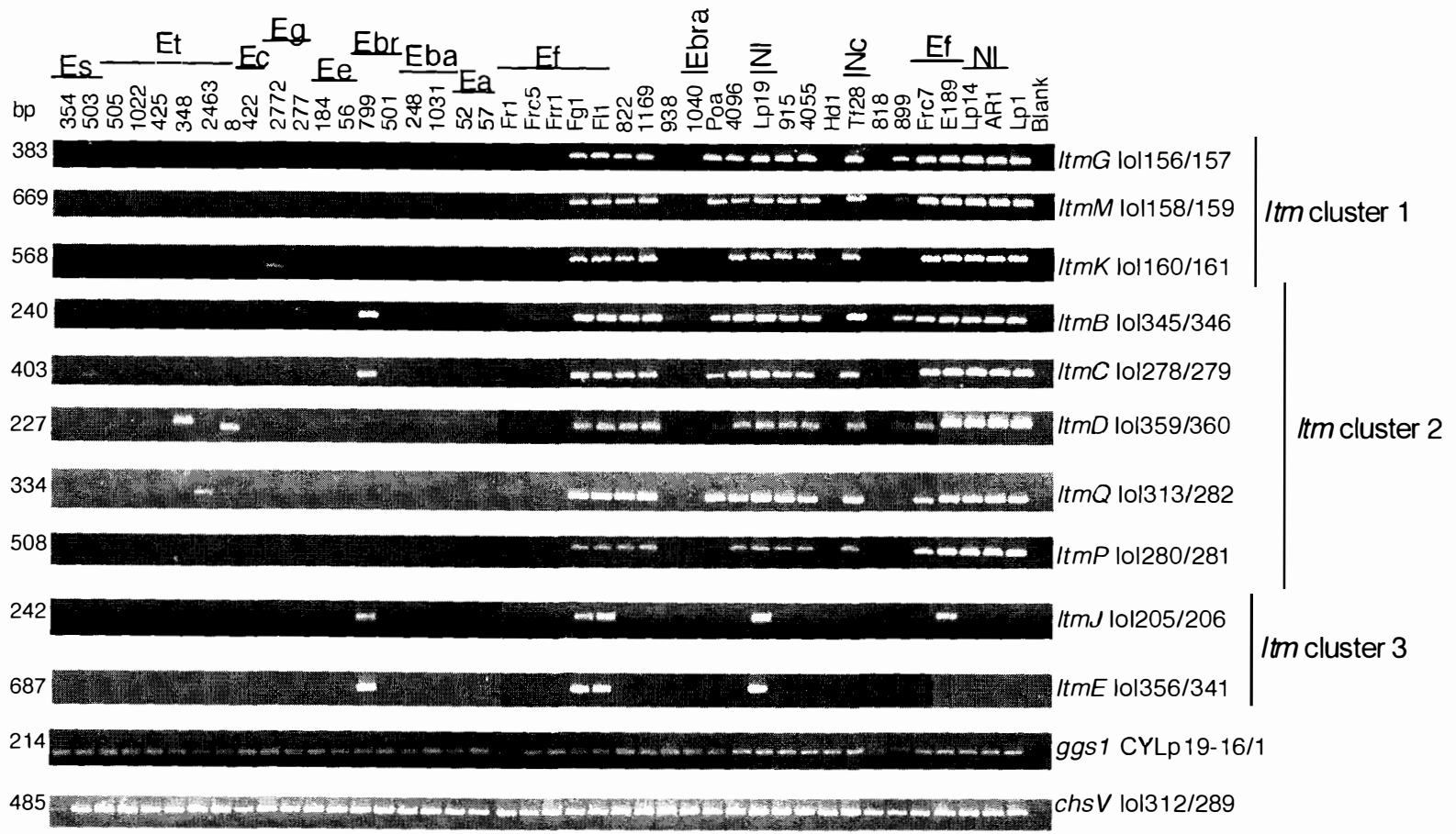


Figure 3.57 PCR screen for the *ltm* genes

Each gene was amplified with primer pairs designed to conserved regions of the gene. The primer pairs used are indicated next to the name of the gene amplified. The primer locations can be found in Appendix 5.1. The size (bp) of the PCR product is on the left margin. Each sample was amplified using the following reaction conditions: 5 ng of genomic DNA, 1x Taq polymerase buffer (Roche), 50 μM each dNTP, 200 nM each primer, 0.5 U Taq polymerase in a reaction volume of 25 μL. The PCR amplification conditions were: 94^{o}C for 2 min followed by 30 cycles of 94^{o}C for 15 s, 58^{o}C for 30 s and 72^{o}C for 1 min; then one cycle of 72^{o}C for 10 min.

The species abbreviations are as follows: ES = *E. sylvatica*; Et = *E. typhina*; Ec = *E. clarkii;* Eg = *E. glyceriae*; Ee = *E. elymi*; Ebr = *E. bromicola*; Eba = *E. baconii*; Ea = *E. amarillans*; Ef = *E. festucae*; Ebra = *E. brachyelytri*; Nl = *N. lolii*; Nc = *N. coenophialum.* Other strains can be identified from Table 2.13.

the *ltmE* primers but have hybridised to this fragment (Fig. 3.56; 3.57). The other *Epichloë* species lacked the *ltm* genes or have remnants of these genes (Fig. 3.57). The *Neotyphodium* species with an *E. festucae* ancestral origin are positive for the *ltm* genes of clusters 1 and 2 (Fig. 3.57). This includes *N. coenophialum* Tf28 (origins of *E. festucae*, ETC, LAE), *N.* sp. LpTG2 Lp1 (origins of *E. festucae*, ETC), *N. melicicola* (origins of *E. festucae*, *N. aotearoae*), *N. siegelii* E915 (origins of *E. festucae*, *E. bromicola*) and *N. tembladerae* E1169 (origins of *E. festucae*, ETC). The exceptions to this are *N. australiense* E938 (ancestral origins of *E. festucae* and *E. typhina* complex), which is *ltm* negative; *N. huerfanum* E4055 (ancestral origins from the *E. typhina* complex and *N. aotearoae*), which is positive for *ltm* clusters 1 and 2; and *N. aotearoae* E899 (*N. aotearoae* ancestral origins), which has some *ltm* genes. PCR on these samples was repeated with fresh DNA dilutions confirming the samples had not been mixed up. *E. bromicola* E799 and *E. typhina* Poa contain multiple remnants of the *ltm* genes. Both strains amplify with the *ltmB* and *ltmC* primers, and E799 also contains *ltmJ* and *ltmE*, whereas amplification with Poa DNA can detect *ltmG*, *ltmM*, *ltmQ* and possibly *ltmD* (Fig. 3.57).

An *ltmG* probe was used to confirm that the *Epichloë* strains negative by PCR were also negative by Southern analysis (Fig. 3.58). This showed that *E. amarillans* strain E52, contains an *ltmG*-hybridising band, which was also detected as a faint PCR product with the *ltmG* primers (Fig. 3.58; 3.57). The hybridisation results of the remaining *Epichloë* strains with *ltmG* (Fig. 3.58) were consistent with the PCR data (Fig. 3.57).

The distribution of the retrotransposons and *pks* gene found to flank the *ltm* clusters was examined further by Southern analysis using a selected range of endophyte strains that cover most *Epichloë* groupings (Fig 3.58). Hybridisation was first performed with a *ggs1* fragment to show that the genomic DNA on the filters was of good quality. The *ggs1* probe hybridises to a 4.5-kb band in the *E. festucae* and *Neotyphodium* species. The other *Epichloë* strains have *ggs1*-hybridising bands of different sizes. Strains that are known to be hybrids were observed to contain two copies of the *ggs1* gene. The retrotransposons Tahi and Rua hybridise to *E. festucae* and *E. baconii* but do not hybridise to other *Epichloë* species. The *E. baconii* strain E1031 and the *Neotyphodium* FaTG-2 isolate Tf15 have a large number of Rua elements but fewer Tahi elements. *E. festucae* E189 and Frc5 and *N. lolii* Lp14 have fewer Tahi and Rua elements than either *E. festucae* Fl1 or *N. lolii* Lp19. The hybridisation of the *pks* pseudogene shows no

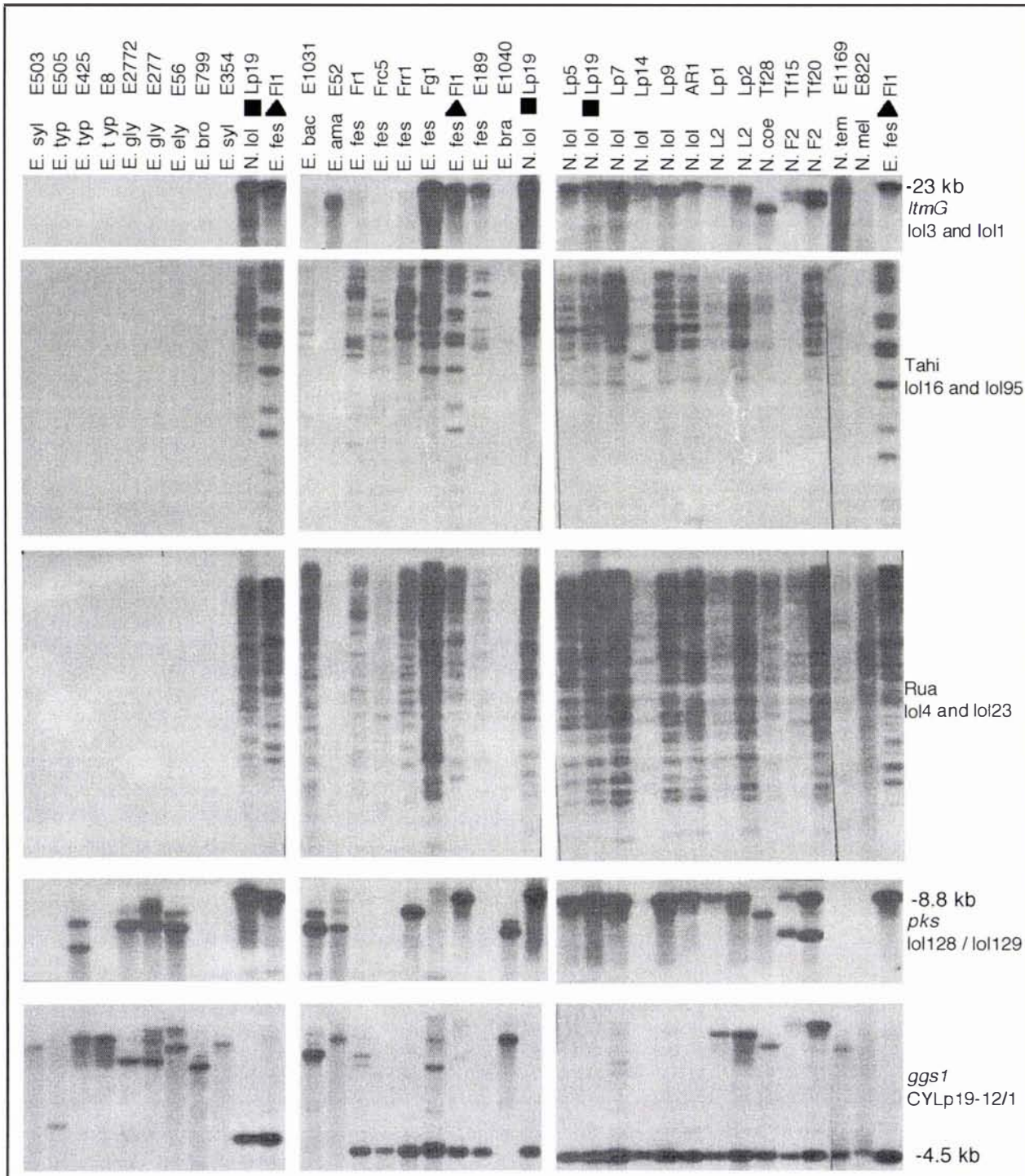


Figure 3.58 Southern analysis on the distribution of the retrotransposons and *pks* sequences

Southern analysis of endophyte strains to determine the presence of the *ltmG*, retrotransposons and *pks* pseudogene. All genomic DNA (2 μg) was digested with *Eco*RI and hybridised with ^{32}P-labelled fragments. The probes and primer pairs used to amplify each fragment are stated at the right of the figure. The strains Lp19 (marked as ■) and Fl1(marked as ▲) have been included on each blot as positive controls and as size markers for each hybridising fragment. All the autoradiographs were over-exposed to detect any faint hybridisation (data not shown). The species abbreviations are as follows; E. ama, *E. amarillans*; E. bac, *E. baconii*; E. bra, *E. brachyelytri*; E. bro, *E. bromicola*; E. ely, *E. elymi*; E. fes, *E. festucae*; E. gly, *E. glyceriae*; E. syl, *E. sylvatica*; E. typ, *E. typhina*; N. F2, *N.* sp. FaTG-2; N. L2, *N.* sp. LpTG-2; N. coe, *N. coenophialum*; N. lol, *N. lolii*; N. mel, *N. melicicola;* N. tem, *N. tembladarae*.

obvious phylogenetic distribution, being present in some strains but not in others. Although the *pks* has been used as a probe for Southern analysis, the *pks* gene from these positive strains have not been sequenced so as yet we do not know if the gene is functional in other endophyte isolates.

A summary of the *ltm* gene distribution based on PCR and Southern analysis and, where known, the lolitrem phenotype is given in Table 3.14. On the basis of these data, functional pathways for lolitrem biosynthesis appear to be confined to the *E. festucae* and *N. lolii* lineage. Remnants of this pathway are found in *E. amarillans*, *E. bromicola* and *E. typhina*. The *ltm* genes are found in many hybrid *Neotyphodium* sp, where nearly all *ltm*-positive strains have an *E. festucae* ancestor. Based on the PCR and Southern data, predictions can be made on the lolitrem phenotype of the endophyte. Strains that lack the essential early pathway genes, such as *ltmG*, *ltmM* and *ltmC*, will not synthesise any indole-diterpene including lolitrems. Such predictions have been summarised in Table 3.14. The data in Table 3.14 indicated that some isolates previously known to be indole diterpene negative, such as *N. lolii* ARI, *N.* spp. Lp1 and *N. coenophialum* Tf28, should be categorised as having the potential to produce indole diterpenes as they contain the essential early pathway genes. These isolates will require further analysis to establish if they are able to produce an earlier intermediate of lolitrem B. Interestingly, *N. lolii* isolates Lp14 and AR1 have identical *ltm* gene distribution patterns yet they have very different indole diterpene phenotypes. This data suggests that Lp14 must contain additional uncharacterised genes that would be required for janthitrem production.

The validity of the PCR screen will require confirmation of the indole diterpene phenotypes by HPLC and mass spectrophotometry methods. However, it does give a clear indication that strains that lack the key early genes would be very unlikely to have the capability to produce indole diterpenes.

3.11 *ltm* homologues detected in other fungi

Genes involved in indole-diterpene production have been cloned and characterised from the filamentous fungi *P. paxilli*, and *A. flavus* (Young, et al., 2001; McMillan, et al., 2003; Zhang, et al., 2004). Genomic organisation of genes shared between the clusters

Table 3.14 *ltm* genotypes of *Epichloë* and *Neotyphodium* endophytes

Species	Isolate	*ltm* genes										*ggs1*	*chsV*	*pks*	Screening technique[1]	Ancestral origin[2]	Predicted ID phenotype[3]	Clay and Schardl[5] 2002
		G	M	K	B	C	D	Q	P	J	E							
E. amarillans	E52	yes										yes	yes	yes	PCR/S		negative	-
E. amarillans	E57											yes	yes		PCR		negative	-
E. baconii	E248											yes	yes		PCR		negative	nt
E. baconii	E1031											yes	yes	yes	PCR/S		negative	nt
E. brachyelytri	E1040				yes							yes	yes	yes	PCR/S		negative	nt
E. bromicola	501											yes	yes		PCR		negative	nt
E. bromicola	E799				yes	yes				yes	yes	yes	yes		PCR/S		negative	nt
E. clarkii	E422											yes	yes		PCR		negative	nt
E. elymi	E56											yes	yes	yes	PCR/S		negative	-
E. elymi	E184											yes	yes		PCR		negative	-
E. festucae	Fr1											yes	yes		PCR/S		negative	+
E. festucae	Frc5											yes	yes		PCR/S		negative	+
E. festucae	Frr1				yes							yes	yes	yes	PCR/S		negative	+
E. festucae	Frc7	yes	yes	yes	yes	yes	yes	yes	yes			yes	yes		PCR/S		positive	+
E. festucae	Fl1	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	PCR/S		positive*	+
E. festucae	Fg1	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes		PCR/S		positive	+
E. festucae	E189	yes	yes	yes	yes	yes	yes	yes	yes	yes		yes	yes		PCR/S		positive	+
E. glyceriae	E2772		yes	yes								yes	yes	yes	PCR/S		negative	nt
E. glyceriae	E277											yes	yes	yes	PCR/S		negative	nt
E. sylvatica	E503					yes	yes					yes	yes		PCR/S		negative	nt
E. sylvatica	E354											yes	yes		PCR/S		negative	nt
E. typhina	E505											yes	yes		PCR/S		negative	-
E. typhina	E425						yes					yes	yes	yes	PCR/S		negative	-
E. typhina	E8						yes					yes	yes		PCR/S		negative*	-
E. typhina	E1022											yes	yes		PCR		negative	-
E. typhina	E348						yes					yes	yes		PCR		negative	-
E. typhina	E2463						yes	yes				yes	yes		PCR		negative	-
E. typhina	Poa	yes	yes		yes	yes	yes	yes	yes			yes	yes		PCR/S	ETC	negative	-
N. aotearoae	E899	yes	yes		yes		yes					yes	yes		PCR	Nao	unsure	-
N. australiense	E938				yes							yes	yes		PCR	Efe, ETC	negative	nt
N. coenophialum	Tf28	yes	yes	yes	yes	yes	yes	yes	yes			yes	yes	yes	PCR/S	Efe, ETC, LAE	negative	-
N. huerfanum	E4055	yes	yes	yes	yes	yes	yes	yes	yes			yes	yes		PCR	ETC	positive	nt
N. inebrians	E818											yes	yes		PCR	Nin	negative	-
N. lolii	Lp19	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	PCR/S	Efe	positive*	+
N. lolii	Lp5	yes	yes	yes	nt	nt	nt	nt	nt	nt	nt	yes	nt	yes	S	Efe	positive*	+
N. lolii	Lp7	yes	yes	yes	nt	nt	nt	nt	nt	nt	nt	yes	nt	yes	S	Efe	positive*	+
N. lolii	Lp14	yes	yes	yes	yes	yes	yes	yes	yes			yes	yes		PCR/S	Efe	positive[4]	+
N. lolii	Lp9	yes	yes	yes	nt	nt	nt	nt	nt	nt	nt	yes	nt	yes	S	Efe	negative*	+
N. lolii	AR1	yes	yes	yes	yes	yes	yes	yes	yes			yes	yes	yes	PCR/S	Efe	negative*	+
N. melicicola	E822	yes	yes	yes	yes	yes	yes	yes	yes			yes	yes		PCR/S	Efe, Nao	positve	-
N. siegelii	E915	yes	yes	yes	yes	yes	yes	yes	yes			yes	yes		PCR/S	Efe, Ebr	positve	-
N. sp. FaTG-2	Tf15	yes	yes	yes	nt	nt	nt	nt	nt	nt	nt	yes	nt	yes	S	Efe, LAE	Positive*	+
N. sp. FaTG-2	Tf20	yes	yes	yes	nt	nt	nt	nt	nt	nt	nt	yes	nt	yes	S	Efe, LAE	unsure	+
N. sp. LpTG-2	Lp1	yes	yes	yes	yes	yes	yes	yes	yes			yes	yes	yes	PCR/S	Efe, ETC	negative*	-
N. sp. LpTG-2	Lp2	yes	yes	yes	nt	nt	nt	nt	nt	nt	nt	yes	nt	yes	S	Efe, ETC	negative*	-
N. sp. "sleepygrass"	E4096	yes	yes	yes	yes	yes	yes	yes	yes			yes	yes		PCR	Efe, Eel	positive	nt
N. tembladarae	E1169	yes	yes	yes	yes	yes	yes	yes	yes			yes	yes		PCR/S	Efe, ETC	positive	nt
N. sp.	Hd1											yes	yes		PCR	ETC, Ebr	negative	nt

[1] Samples were screened by PCR only (PCR), Southern analysis only (S) or both Southern analysis and PCR (PCR/S).

[2] Ancestral origin of the asexual species as defined by Moon et al 2004

[3] The indole-diterpene phenotype was predicted based on the presence of the ltm genes, or known (*).
An indole-diterpene positive phenotype is given to strains that are positive for *ltmG*, *ltmM* and *ltmC*.
Where the results are incomplete a label of unsure has been used.
Known positive phenotypes are highlighted in green, known negative phenotypes are highlighted in red.

[4] Lp14 is known to make indole-diterpenes of the janthitrem group but does not produce lolitrems.

[5] Indole-ditepene phenotype as summarised in Clay and Schardl, 2002. '-' = undetected from any isolate tested, '+' detected from some isolates
nt = not tested
yes highlighted in pale green are samples that have a band by PCR but no sequence analysis to confirm it is a related gene

of *P. paxilli*, *A. flavus* and *N. lolii* strains are shown in Figure 3.59. Orthologues of *ltmG*, *ltmM* and *ltmC* were found in both the *pax* and *atm* clusters. The *ltm* cluster also contained homologues of *paxB*, *paxD*, *paxP* and *paxQ*. The *atm* cluster is only partially sequenced, but based on the structural similarity of paxilline to aflatrem, the *atm* cluster was predicted to contain other *pax* homologues such as *paxP*, *paxQ*, *paxD* and possibly *paxB* (Zhang, et al., 2004). Sequence comparisons of the *ltm* genes to the *pax* and *atm* genes are summarised in Table 3.12 and alignments are shown in Appendix 5.1. The location and order of the genes from the *pax*, *ltm* and *atm* clusters were different among the strains (Fig. 3.59). As expected, the *ltm* cluster contained additional genes with putative roles for the decoration of the indole-diterpene backbone.

A database search of the four 'completed' fungal genomes of *N. crassa*, *A. nidulans*, *M. grisea* and *F. graminearum* was used to identify *ltm* homologues from these species. BLASTP analysis was used to find *ltm* homologues and their genomic location (Table 3.15). Sequences contained within the four fungal genomes, with BLASTP matches to *ltm* genes that were better than 1e-10, were analysed for linkage of two or more *ltm* homologues.

The *N. crassa* genome contains homologues of *ltmG*, *ltmM*, *ltmD*, *ltmB*, *ltmQ* and *ltmK*. However, there is only one copy of geranylgeranyl diphosphate synthase, which is the *al-3* gene (Barbato, et al., 1996), a *ggs1*-type rather than an *ltmG*. The remaining genes have BLASTP E-values greater than 1e-50 and none of these genes was located in clusters.

The *F. graminearum* genome contained sequence homologues of *ltmG*, *ltmM*, *ltmB* and the P450 monooxygenases but did not contain an *ltmD* (*dmaW*-like prenyl transferase) homologue. The BLASTP matches to the P450 monooxygenase sequences all have E-values greater than 4e-52. There are three putative *ggs* genes, the *ggs1* (FG10097, Section 3.2.1), a second *ggs*, FG04591, clustered with *ltmB* and *ltmM* homologues and a third, FG01738, with a less significant BLASTP match (Table 3.15).

The *M. grisea* genome contains homologues of the *ltm* genes but the BLASTP values are greater than 1e-51. There were five regions that contained clustered homologues of the *ltm* genes with the most significant cluster containing a cytochrome P450 monooxygenase, a prenyl transferase and a FAD-dependent monooxygenase (Table

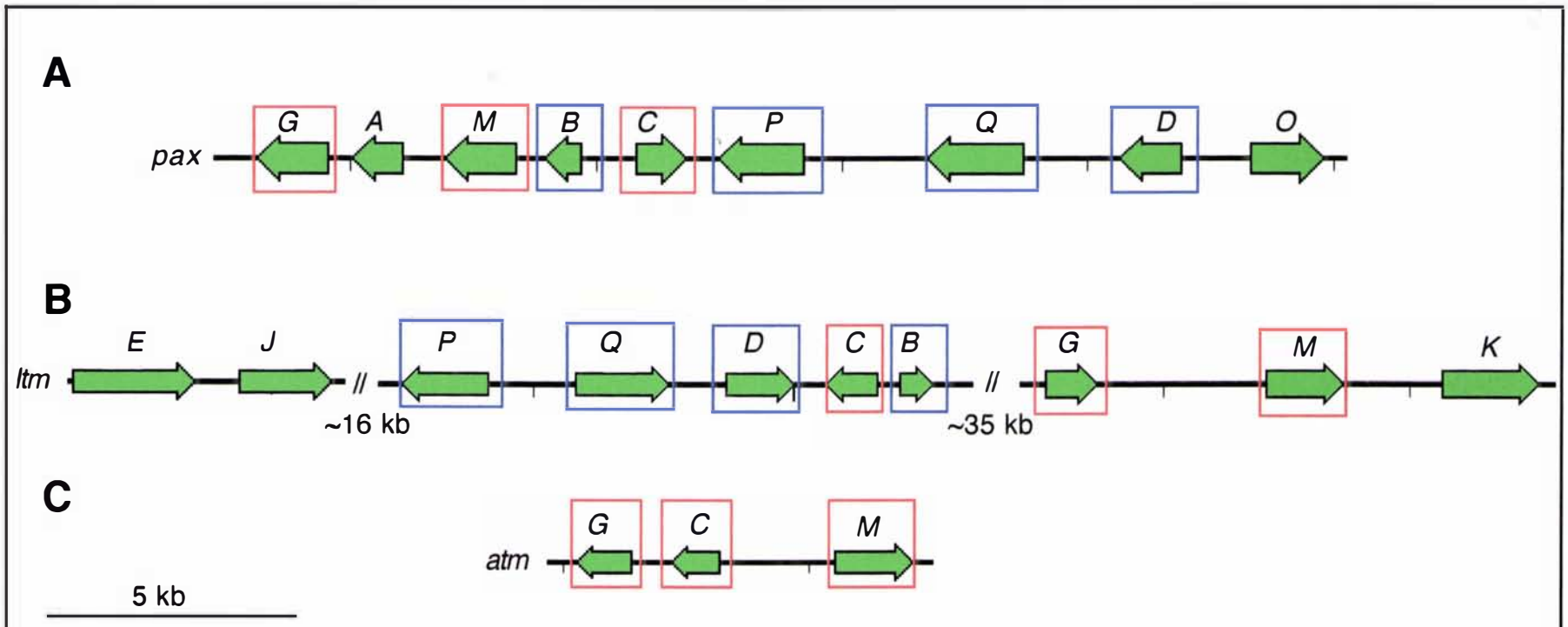


Figure 3.59 Comparison of indole-diterpene gene clusters from *P. paxilli*, *N. lolii* and *A. flavus*

Genomic organisation of the (**A**) *P. paxilli pax*, (**B**) *N. lolii ltm* and (**C**) *A. flavus atm* gene clusters. The arrows indicate the position and direction of transcription of each gene. The gene names have been abbreviated to a single code letter. The genes with red boxes have been identified in all three clusters. The genes in blue boxes have been identified in the *pax* and *ltm* gene clusters. However, the sequence of the *atm* cluster is incomplete.

Table 3.15 Homologues of *ltm* genes contained in *N. crassa*, *A. nidulans*, *M. grisea*, and *F. graminearum*

Gene ID numer[1]	Species	Contig[2]	Putative function[3]	*ltm*[4]	E-value[5]	*ltm*[4]	E-value[5]	*ltm*[4]	E-value	*ltm*[4]	E-value
NCU01427	*N. crassa*	3.57	al-3, geranylgeranyl diphosphate synthase	G	7e-92	E	1e-18	C	1e-10		
NCU02925	*N. crassa*	3.152	FAD monooxygenase	M	7e-25						
NCU04693	*N. crassa*	3.257	prenyl transferase dmaW-like	E	1e-34	D	2e-33				
NCU05376	*N. crassa*	3.306	Cytochrome P450 monooxygenase	Q	2e-36	K	2e-35	P	1e-30	J	1e-29
NCU05967	*N. crassa*	3.339	Cytochrome P450 monooxygenase	Q	2e-26	J	7e-23	K	2e-23		
NCU08173	*N. crassa*	3.492	FAD monooxygenase	M	4e-15						
NCU09274	*N. crassa*	3.580	Cytochrome P450 monooxygenase	K	4e-48	J	2e-43	P	3e-41	Q	5e-22
NCU09400	*N. crassa*	3.598	unknown	B	7e-20						
AN0654	*A. nidulans*	1.10	geranylgeranyl diphosphate synthase	G	6e-89	E	4e-18	C	8e-12		
AN1592	*A. nidulans*	1.25	prenyl transferase	G	3e-69	E	2e-17				
AN1598	*A. nidulans*	1.25	Cytochrome P450 monooxygenase	K	3e-84	J	2e-81	P	4e-71	Q	3e-46
AN1784	*A. nidulans*	1.28	FAD monooxygenase	M	9e-68						
AN2114	*A. nidulans*	1.32	FAD monooxygenase	M	5e-15						
AN2345	*A. nidulans*	1.38	prenyl transferase dmaW-like	E	1e-21	D	5e-19				
AN2407	*A. nidulans*	1.39	prenyl transferase	G	1e-40	E	2e-10				
AN2569	*A. nidulans*	1.43	Cytochrome P450 monooxygenase	J	9e-16	P	9e-15	K	1e-14		
AN2611	*A. nidulans*	1.43	prenyl transferase ggs-like	G	8e-35	E	5e-14				
AN3253	*A. nidulans*	1.54	Cytochrome P450 monooxygenase	K	5e-53	J	8e-46	P	2e-45	Q	1e-27
AN3256	*A. nidulans*	1.54	Cytochrome P450 monooxygenase	K	5e-43	J	2e-38	P	1e-31	Q	1e-27
AN3274	*A. nidulans*	1.55	Cytochrome P450 monooxygenase	K	2e-48	J	8e-44	P	1e-42	Q	6e-41
AN3497	*A. nidulans*	1.59	Cytochrome P450 monooxygenase	Q	5e-30	J	4e-23	P	7e-20	K	2e-16
AN3596	*A. nidulans*	1.61	FAD monooxygenase	M	4e-13						
AN4576	*A. nidulans*	1.78	FAD monooxygenase	M	5e-10						
AN6018	*A. nidulans*	1.103	FAD monooxygenase	M	6e-46						
AN6784	*A. nidulans*	1.112	prenyl transferase	E	5e-16						
AN6810	*A. nidulans*	1.113	prenyl transferase	G	6e-59	E	2e-17	C	1e-15		
AN7684	*A. nidulans*	1.130	FAD monooxygenase	M	3e-11						
AN8139	*A. nidulans*	1.141	Cytochrome P450 monooxygenase	Q	9e-30	K	8e-30	P	4e-23	J	3e-22
AN8140	*A. nidulans*	1.141	FAD monooxygenase	M	2e-43						
AN8141	*A. nidulans*	1.141	Cytochrome P450 monooxygenase	P	2e-31	Q	5e-27	K	1e-23	J	3e-20
AN8143	*A. nidulans*	1.141	prenyl transferase ggs-like	G	2e-81	E	7e-13				
AN8144	*A. nidulans*	1.141	unknown	B	1e-25						
AN8184	*A. nidulans*	1.141	Cytochrome P450 monooxygenase	P	4e-63	K	3e-54	J	1e-39	Q	1e-38
AN8250	*A. nidulans*	1.145	Cytochrome P450 monooxygenase	J	9e-24	P	1e-21	Q	1e-19	K	2e-17
AN8378	*A. nidulans*	1.152	FAD monooxygenase	M	4e-67						
AN8470	*A. nidulans*	1.153	FAD monooxygenase	M	8e-13						
AN8482	*A. nidulans*	1.153	prenyl transferase	E	6e-19						
AN8514	*A. nidulans*	1.154	prenyl transferase dmaW-like	E	2e-58	D	2e-49				
AN8530	*A. nidulans*	1.155	Cytochrome P450 monooxygenase	K	2e-37	P	6e-32	J	2e-29	Q	4e-27
AN9229	*A. nidulans*	1.170	prenyl transferase	E	5e-11						
AN9230	*A. nidulans*	1.170	FAD monooxygenase	M	2e-36						
AN9248	*A. nidulans*	1.172	Cytochrome P450 monooxygenase	K	2e-32	P	2e-31	Q	2e-30	J	1e-26
AN9251	*A. nidulans*	1.172	Cytochrome P450 monooxygenase	J	2e-14	K	8e-12	P	1e-12		
AN9253	*A. nidulans*	1.172	Cytochrome P450 monooxygenase	P	2e-31	Q	5e-27	K	1e-23	J	3e-20
AN9257	*A. nidulans*	1.172	unknown	B	2e-19						
AN9258	*A. nidulans*	1.172	FAD monooxygenase	M	6e-80						
FG00057	*F. graminearum*	1.2	FAD monooxygenase	M	1e-49						
FG00071	*F. graminearum*	1.4	cytochrome P450 monooxygenase	K	5e-33	Q	2e-33	J	1e-30	P	5e-27
FG00092	*F. graminearum*	1.5	FAD monooxygenase	M	1e-15						
FG01738	*F. graminearum*	1.92	prenyl transferase ggs-like	G	4e-51	E	3e-17	C	5e-12		
FG01786	*F. graminearum*	1.93	cytochrome P450 monooxygenase	K	4e-16	J	8e-12				
FG01819	*F. graminearum*	1.93	FAD monooxygenase	M	4e-45						
FG01868	*F. graminearum*	1.100	cytochrome P450 monooxygenase	K	4e-52	J	1e-41	P	1e-37	Q	2e-31
FG02117	*F. graminearum*	1.111	cytochrome P450 monooxygenase	J	8e-22	K	5e-22	Q	2e-22	P	9e-21
FG02672	*F. graminearum*	1.132	cytochrome P450 monooxygenase	Q	7e-35	J	2e-32	K	2e-31	P	8e-30
FG03657	*F. graminearum*	1.160	FAD monooxygenase	M	8e-11						
FG03729	*F. graminearum*	1.162	FAD monooxygenase	M	2e-12						
FG03775	*F. graminearum*	1.163	FAD monooxygenase	M	5e-26						
FG04461	*F. graminearum*	1.192	unknown	B	2e-18						
FG04591	*F. graminearum*	1.193	prenyl transferase ggs-like	G	2e-94	E	2e-16	C	9e-10		
FG04594	*F. graminearum*	1.193	Unknown	B	2e-51						
FG04595	*F. graminearum*	1.193	FAD monooxygenase	M	3e-63						
FG04717	*F. graminearum*	1.196	cytochrome P450 monooxygenase	K	4e-34	Q	1e-30	P	7e-28	J	3e-28
FG04776	*F. graminearum*	1.196	FAD monooxygenase	M	4e-12						
FG05063	*F. graminearum*	1.203	FAD monooxygenase	M	2e-15						
FG05334	*F. graminearum*	1.216	cytochrome P450 monooxygenase	Q	7e-16	K	9e-13	J	3e-10		
FG08116	*F. graminearum*	1.325	FAD monooxygenase	M	1e-14						
FG08187	*F. graminearum*	1.329	cytochrome P450 monooxygenase	Q	9e-30	P	4e-28	K	7e-27	J	4e-24
FG08228	*F. graminearum*	1.329	FAD monooxygenase	M	3e-11						
FG09063	*F. graminearum*	1.370	FAD monooxygenase	M	5e-10						
FG09195	*F. graminearum*	1.371	cytochrome P450 monooxygenase	J	4e-25	P	2e-25	K	8e-20	Q	5e-20
FG10097	*F. graminearum*	1.418	geranylgeranyl diphosphate synthase	G	3e-90	E	1e-21	C	3e-13		
FG10612	*F. graminearum*	1.444	FAD monooxygenase	M	9e-15						
FG11002	*F. graminearum*	1.457	cytochrome P450 monooxygenase	Q	2e-27	P	5e-26	K	5e-24	J	1e-23
FG11100	*F. graminearum*	1.459	FAD monooxygenase	M	7e-30						
FG11215	*F. graminearum*	1.460	FAD monooxygenase	M	8e-56						
FG11282	*F. graminearum*	1.464	cytochrome P450 monooxygenase	K	9e-40	Q	5e-38	J	3e-35	P	1e-29
FG11536	*F. graminearum*	1.473	cytochrome P450 monooxygenase	K	5e-41	P	2e-41	J	3e-40	Q	5e-31
MG00024	*M. grisea*	2.7	cytochrome P450 monooxygenase	J	2e-33	Q	3e-31	P	6e-30	K	1e-30
MG00026	*M. grisea*	2.7	prenyl transferase ggs-like	G	2e-35	E	5e-14	C	6e-10		
MG00031	*M. grisea*	2.7	FAD monooxygenase	M	1e-51						
MG00238	*M. grisea*	2.57	cytochrome P450 monooxygenase	K	4e-29	P	3e-28	J	2e-25	Q	9e-24
MG00300	*M. grisea*	2.70	cytochrome P450 monooxygenase	J	6e-31	Q	3e-31	K	4e-26	P	2e-17
MG00572	*M. grisea*	2.99	cytochrome P450 monooxygenase	K	5e-29	Q	1e-25	P	3e-22	J	2e-20
MG00758	*M. grisea*	2.134	geranylgeranyl diphosphate synthase	G	1e-88	E	1e-21	C	5e-11		
MG01391	*M. grisea*	2.256	cytochrome P450 monooxygenase	K	1e-36	P	8e-35	J	3e-33	Q	9e-32
MG01700	*M. grisea*	2.318	cytochrome P450 monooxygenase	K	6e-34	Q	2e-28	P	6e-24	J	1e-22

Gene ID numer[1]	*Species*	Contig[2]	Putative function[3]	*ltm*[4]	E-value[5]	*ltm*[4]	E-value[5]	*ltm*[4]	E-value	*ltm*[4]	E-value
MG01701	*M. grisea*	2.318	prenyl transferase ggs-like	G	4e-42						
MG01950	*M. grisea*	2.368	cytochrome P450 monooxygenase	K	1e-41	P	5e-39	J	2e-33	Q	4e-27
MG02256	*M. grisea*	2.442	FAD monooxygenase	M	1e-35						
MG02272	*M. grisea*	2.445	cytochrome P450 monooxygenase	J	9e-48	K	2e-47	Q	2e-30	P	3e-26
MG02292	*M. grisea*	2.449	FAD monooxygenase	M	2e-24						
MG02294	*M. grisea*	2.449	cytochrome P450 monooxygenase	J	2e-33	K	1e-32	P	9e-31	Q	2e-31
MG03104	*M. grisea*	2.617	cytochrome P450 monooxygenase	K	7e-18	Q	6e-17	J	6e-15	P	2e-15
MG03432	*M. grisea*	2.675	prenyl transferase ggs-like	G	6e-43	E	6e-12				
MG03834	*M. grisea*	2.748	cytochrome P450 monooxygenase	K	4e-36	J	9e-26	Q	9e-24	P	1e-24
MG04349	*M. grisea*	2.822	cytochrome P450 monooxygenase	J	6e-22	P	2e-21	Q	1e-21	K	6e-18
MG06540	*M. grisea*	2.1218	prenyl transferase dmaW-like	D	3e-40	E	2e-32				
MG06761	*M. grisea*	2.1252	FAD monooxygenase	M	8e-19						
MG07247	*M. grisea*	2.1347	cytochrome P450 monooxygenase	J	3e-21	Q	5e-20	K	2e-17		
MG07506	*M. grisea*	2.1395	prenyl transferase ggs-like	G	8e-41	E	1e-13				
MG08280	*M. grisea*	2.1551	prenyl transferase dmaW-like	D	2e-30	E	1e-22				
MG08293	*M. grisea*	2.1557	FAD monooxygenase	M	6e-11						
MG08378	*M. grisea*	2.1577	cytochrome P450 monooxygenase	K	2e-23	J	6e-20	Q	1e-20	P	5e-19
MG08379	*M. grisea*	2.1577	cytochrome P450 monooxygenase	K	3e-30	P	7e-25	Q	6e-23	J	5e-17
MG08385	*M. grisea*	2.1578	cytochrome P450 monooxygenase	K	8e-28	Q	1e-26	J	3e-23	P	1e-20
MG08387	*M. grisea*	2.1579	cytochrome P450 monooxygenase	J	2e-24	K	2e-23	Q	6e-17	P	4e-13
MG08459	*M. grisea*	2.1594	cytochrome P450 monooxygenase	P	3e-39	J	7e-37	K	2e-36	Q	3e-26
MG08671	*M. grisea*	2.1638	unknown	B	1e-14						
MG09156	*M. grisea*	2.1746	cytochrome P450 monooxygenase	Q	9e-32	J	3e-20	P	3e-19	K	5e-14
MG09198	*M. grisea*	2.1756	cytochrome P450 monooxygenase	K	6e-36	J	7e-31	P	4e-28	Q	1e-23
MG09448	*M. grisea*	2.1810	prenyl transferase ggs-like	G	2e-38	E	1e-19	C	2e-13		
MG09453	*M. grisea*	2.1810	FAD monooxygenase	M	9e-41						
MG10014	*M. grisea*	2.1916	FAD monooxygenase	M	6e-10						
MG10345	*M. grisea*	2.1991	cytochrome P450 monooxygenase	Q	9e-27	K	4e-27	P	1e-26	J	2e-21
MG10351	*M. grisea*	2.1991	cytochrome P450 monooxygenase	K	1e-37	J	9e-34	P	2e-30	Q	2e-28
MG10527	*M. grisea*	2.2024	cytochrome P450 monooxygenase	K	3e-34	J	1e-32	Q	9e-29	P	1e-28
MG10746	*M. grisea*	2.2065	unknown	B	1e-15						
MG10908	*M. grisea*	2.2111	FAD monooxygenase	M	2e-27						
MG10953	*M. grisea*	2.2124	prenyl transferase dmaW-like	E	3e-16	D	3e-10				
MG11075	*M. grisea*	2.2223	cytochrome P450 monooxygenase	P	1e-34	J	4e-33	K	2e-32	Q	2e-30

[1] Genes that are clustered or in close proximity are identified in red.

[2] The contig the genes are contained on as defined by the Broad Institute (http://www.broad.mit.edu/annotation/funqi/fqi).

[3] The predicted function of the genes based on homology to the *ltm* genes.

[4] The *ltm* genes with the best BLASTP match. The cytochrome P450 monooxygenases and the prenyl transferases have multiple sequence matches.

[5] E-values highlighed yellow are matches to the *qqs1* genes. E-values highlighted green are greater than 1e-50.

3.15). Associated with each of the *M. grisea* clusters were transposon sequences. There were six putative homologues of prenyl transferases of the *ggs*-type one of which was the *ggs1* gene (MG00758, Section 3.2.1), three were associated with other *ltm* homologues and two were not associated with *ltm* genes (Table 3.15).

The *A. nidulans* genome contained *ltm* homologues with significant BLASTP matches (E-values less than 1e-49) to *ltmG*, *ltmK*, *ltmM*, *ltmP*, and *ltmD* (Table 3.15). There were two *ggs* homologues, AN0654 (a *ggs1*, Section 3.2.1) and a second homologue, AN8143, which is associated within a cluster that contained *ltmB*, *ltmM*, *ltmP* and *ltmQ* homologues (Table 3.15). There were two additional *ggs*-like prenyl transferases with BLASTP matches that were less significant than those mentioned above, of which one (AN1592) was clustered with an *ltmK* homologue (Table 3.15). The *A. nidulans* genome contained four regions with clustered *ltm* homologues of which three of these had genes within the clusters that had highly significant BLASTP matches (E-value less than 1e-69) to *ltm* genes (Table 3.15).

To determine the phylogenetic relationship between the *ggs1*-like and those required for secondary metabolism (*ggs2*-like) the fungal GGPP synthase sequences from *N. lolii*, *E. festucae*, *P. paxilli*, *F. fujikuroi*, *A. flavus*, *F. graminearum*, *N. crassa*, *M. grisea*, *E. typhina*, *A. nidulans* were aligned using ClustalX (Appendix 5.1). The phylogenetic relationship of the GGPP synthases was shown in an unrooted tree based on a split decomposition graph, which was further supported by the Neighbour-Joining method (Fig. 3.60). The GGPP synthase sequences were separated into two distinct groups, containing the *ggs1*-like and the *ggs2*-like sequences, respectively. Each group contained sequences that were more similar between species than within their own species. Each *ggs1*-like sequence contained the IPPRXSS motif mentioned in Section 3.2.1 (Appendix 5.1).

The phylogenetic tree supports the hypothesis that *F. graminearum* and *A. nidulans* have two copies of the *ggs* gene, one that clusters with the house keeping *ggs1*-like and a second gene that clusters with the *ggs2*-like required for secondary metabolism. To date there are no reports of *A. nidulans*, *M. grisea* or *F. graminearum* producing diterpene compounds.

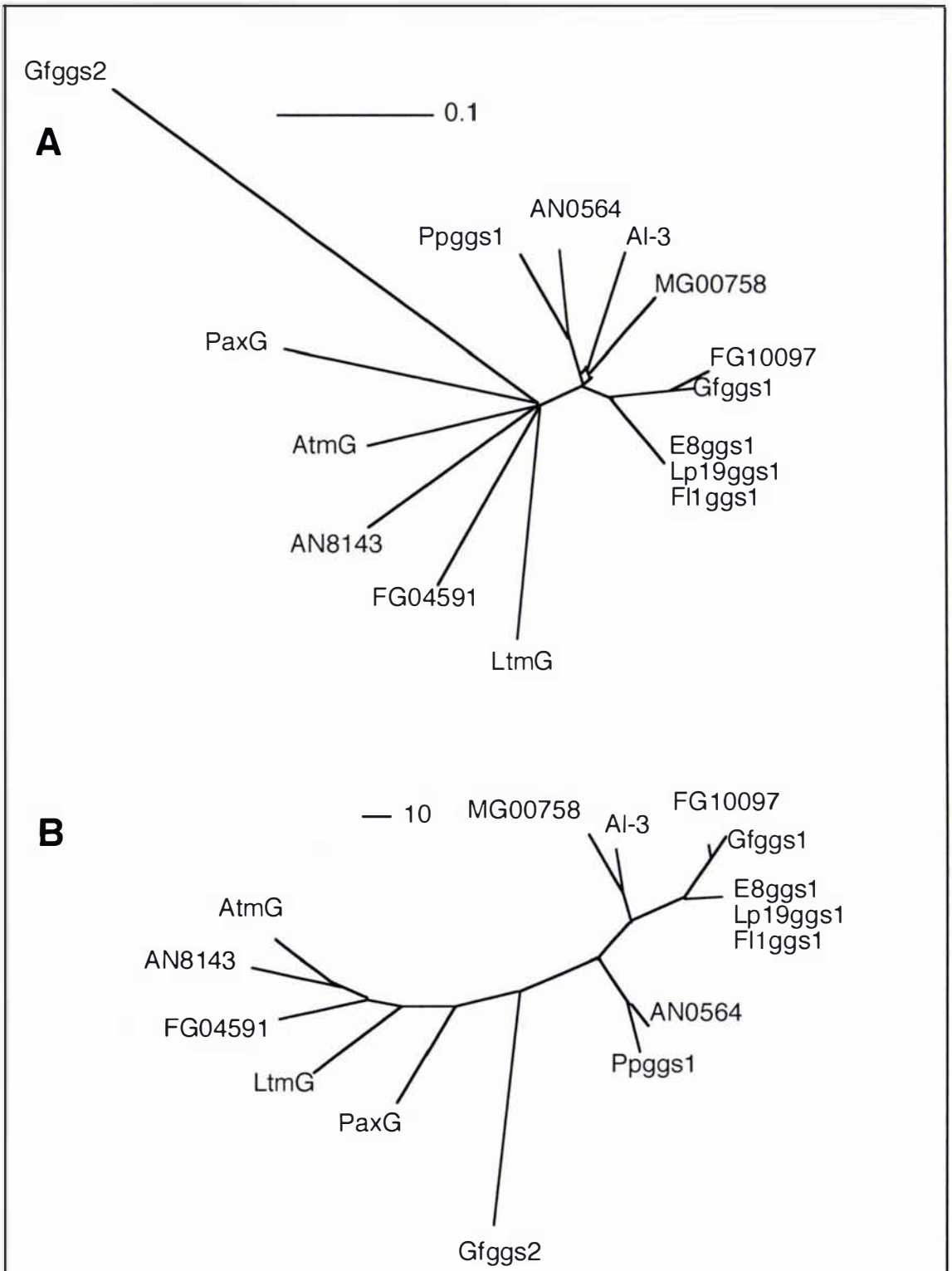


Figure 3.60 Phylogenetic analysis of the fungal GGPP synthases

(A) A phylogenetic tree based on a split decomposition graph. **(B)** A phylogenetic tree based on Neighbour Joining graph. Both Trees were generated by Pete Lockhart. The GGPP synthase sequences are as follows *P. paxilli* (PaxG accession number AAK11531; Ppggs1 accession number AAK11525), *A. nidulans* (AN0564 accession number EAA65430; AN8143 accession number EAA59165), *A. flavus* (AtmG accession number AAT65717), *G. fujikuroi* (Gfggs1 accession number EAA68323; Gfggs2 accession number CAA75568), *F. graminearum* (FG04591 accession number EAA72205; FG10097 accession number EAA68323), *N. crassa* (Al-3 accession number CAD70868), *M. grisea* (MG00758 accession number EAA49100), *N. lolii* Lp19ggs1 and LtmG, *E. festucae* Fl1ggs1 and *E. typhina* E8ggs1.

CHAPTER FOUR

DISCUSSION

4.1 Endophytes that produce lolitrem B contain two *ggs* genes

Penicillium paxilli produces the indole-diterpene alkaloid paxilline, which has structural similarities to lolitrem B, a secondary metabolite of *Neotyphodium lolii*. Paxilline has been identified in extracts of *N. lolii*-infected perennial ryegrass suggesting that paxilline and/or precursors of paxilline are likely intermediates of lolitrem B biosynthesis (Munday-Finch, et al., 1998; Munday-Finch, et al., 1996a; Mantle, Weedon, 1994). The *P. paxilli pax* biosynthesis gene cluster contains a set of at least five genes that are essential for paxilline production: *paxG* encoding a GGPP synthase, *paxM* encoding an FAD-dependent monooxygenase, *paxC* encoding a prenyl transferase, and *paxP* and *paxQ* encoding two cytochrome P450 monooxygenases (Young, et al., 2001; McMillan, et al., 2003). Within the *pax* cluster, *paxG* is the most highly conserved gene showing significant similarity to available GGPP synthase sequences. In fact, the *P. paxilli* genome contains two independent GGPP synthase genes, *ggs1* and *paxG*. The *paxG* gene was shown by deletion analysis to be pathway specific for paxilline production. The inability of *ggs1* to complement the *paxG* mutation suggests that GGS1 and PaxG are spatially separate in the cell (Young, et al., 2001).

The conserved gene structure of GGPP synthases in filamentous fungi allowed degenerate primers to be designed to isolate the corresponding genes from *N. lolii*. This approach led to the isolation of two GGPP synthases, *ggs1* and *ltmG* (Section 3.1). The *ltmG* gene was only present in the lolitrem-producing isolates, *N. lolii* Lp19 and *E. festucae* Fl1, whereas *ggs1* was identified in both lolitrem-producing and non-producing isolates (Section 3.1). Therefore, *ltmG* is predicted to be the *paxG* orthologue. The presence of two GGPP synthase genes is consistent with data obtained from *P. paxilli* and other filamentous fungi that synthesise diterpenes (Young, et al., 2001; Tudzynski, Hölter, 1998; Zhang, et al., 2004). The presence of two GGPP synthases is now considered a signature of fungal secondary metabolite diterpene production (Parker, Scott, 2004).

The *ggs1* genes from *N. lolii*, *E. festucae* and *E. typhina* lack introns (Section 3.2). Sequence analysis of these *ggs1* genes showed that, although they are highly conserved, there are small polymorphic regions contained within repeated sequences of the *ggs1*

genes. Of particular interest is the polymorphic CAG repeat, which was shown to be useful for strain identification (Section 3.2.2). *Epichloë* and *Neotyphodium* isolates were distinguished based on the number of CAG and flanking CAT repeats. Strains that group within the *Epichloë typhina* complex (ETC), such as *E. typhina* and *E. sylvatica*, contain CAS (where S = G or C) instead of CAG. The sequence variation of the CAG repeat is easily detected by sequencing a PCR product amplified with primers that are specific to all isolates of the *Epichloë* and *Neotyphodium* groups. This simple diagnostic test is a useful tool for distinguishing a large number of isolates and differentiating between closely related *N. lolii* isolates, such as Lp19 and AR1.

The *ltmG* genes contained in the lolitrem B-producing strains *N. lolii* Lp19 and *E. festucae* Fl1 were shown by RT-PCR to contain two introns (Section 3.3) of which the positions and phases are conserved with the first two introns of the *P. paxilli paxG*, *A. flavus atmG* and *F. fujikuroi ggs2* from the paxilline, aflatrem and gibberellin biosynthesis clusters, respectively (Young, et al., 2001; Tudzynski, Hölter, 1998; Zhang, et al., 2004). The expression profiles of *ltmG* and *ggs1* differ. Expression of the *ltmG* gene was not detected in culture but was highly expressed *in planta* (Section 3.5 and 3.9.3), whereas expression of the *ggs1* gene was detected in culture (Section 3.2.3), but was not detected *in planta*. The endophyte DNA biomass from endophyte-infected plant material was estimated to be between 0.3 and 1.9% of the total plant DNA biomass and when this is taken into consideration *ltmG* was shown to be highly expressed *in planta*.

GGPP synthases contain five highly conserved regions (Chen, et al., 1994) of which the first domain in filamentous fungi, is situated approximately 70 residues from the N-terminus. The N-terminal region of GGPP synthases prior to the first conserved region contains the greatest sequence diversity. However, alignment of available fungal GGPP synthase sequences showed that the *ggs1*-like genes have a unique protein motif, IPPRTSS, which is absent from the fungal genes involved in secondary metabolism pathways. The IPPRTSS motif was located within the first 50 amino acids of the *P. paxilli*, *F. fujikuroi*, *E. festucae*, *E. typhina*, *N. lolii*, *A. nidulans*, *M. grisea*, *F. graminearum* and *N. crassa ggs1* genes (Appendix 5.1.1). The IPPRTSS motif was identified in 100 protein sequences, with at most one mismatch, using the BLASTP algorithm to search for short, nearly identical sequences. Nearly fifty percent of the detected sequences were of fungal origin, which indicated that this motif could have a

significant function within fungal proteins. As yet, the IPPRTSS motif has not been characterised in other proteins so it remains uncertain what the likely function of this sequence is, but a putative role could be to anchor or translocate Ggs1 into a subcellular localisation separate from PaxG. Analysis of Ggs1 using the web-based localisation predictions (Section 2.7) did not reveal conclusively the putative subcellular localisation of this protein. However, translocation of GGPP synthases has been shown in *Arabidopsis thaliana*, which contains five independent GGPP synthases. The N-terminal regions of these proteins have been shown to be redundant for GGPP synthase activity. Subsequent fusion of these N-terminal regions to GFP showed that the GFP protein could be translocated to separate subcellular localisations (Okada, et al., 2000). Other possibilities that could limit GGPP synthase complementation are low-level *ggs1* expression, tissue-specific expression of the *ggs2*-like genes and/or protein stability.

Phylogenetic analysis of the fungal GGPP synthases groups the *ggs1*-like housekeeping GGPP synthases of different species together and separate from their homologues, the *ggs2*-like GGPP synthases, which are involved in secondary metabolite biosynthesis (Section 3.11). Two independent *ggs* genes have been identified within the genomes of *F. graminearum* and *A. nidulans*. The diterpenoid phenotypes of these strains are unknown but the phylogenetic analysis established that one of the two genes groups with the *ggs1*-like genes and the other groups with the *ggs2*-like genes. Therefore, if the presence of two GGPP synthases is a signature for diterpene production (Parker, Scott, 2004), then *F. graminearum* and *A. nidulans* would be predicted to synthesise compounds with a diterpene moiety, assuming the genes are functional.

4.2 Identification of 10 *ltm* genes

The indole-diterpene biosynthesis pathway for lolitrem B production is predicted to contain at least 8-12 genes, of which five have been characterised in the paxilline biosynthesis pathway of *P. paxilli* (Young, et al., 2001; McMillan, et al., 2003; Scott, McMillan, Astin, Saikia, Young, Bryant and Parker unpublished results). The first *N. lolii ltm* gene cluster identified within the Lp19 genome contained *ltmG*, *ltmM*, a monooxygenase, and *ltmK*, a P450 monooxygenase with predicted roles in indole-diterpene biosynthesis (Section 3.4). However, flanking these three genes are degenerate retrotransposons and a polyketide synthase pseudogene. The sequence of

ltm cluster 1 is under-represented in the genomic library and the repetitive nature of the retrotransposons and flanking sequences made it difficult to isolate clones in order to sequence these regions (Section 3.7). Three additional genes encoding, a prenyl transferase, ltmC, and two P450 monooxygenases, *ltmP* and *ltmJ*, were subsequently identified within EST libraries, which enabled the isolation of the corresponding genomic regions. The *ltm* cluster 2 was identified by hybridisation with fragments of *ltmC* and *ltmP* and contained 5 genes *ltmC, ltmD, ltmP, ltmQ and ltmB* encoding a prenyl transferase, a dimethylallyl tryptophan (DMAT)-like synthase (prenyl transferase), two P450 monooxygenases, and a protein of unknown function, respectively (Section 3.9.1). The *ltm* cluster 3 was identified by hybridisation with a fragment of *ltmJ* and contains at least 2 genes, *ltmJ* and *ltmE*, encoding a P450 monooxygenase, and a multifunctional enzyme containing a DMAT-like synthase and a PaxC-like prenyl transferase (Section 3.9.2).

The *ltm* biosynthesis gene clusters contain four P450 monooxygenases. Two of these genes, *ltmP* and *ltmQ*, have significant similarity to *paxP* and *paxQ*, respectively, from the *P. paxilli pax* cluster. Based on their significant sequence similarity to *pax* genes, *N. lolii* genes *ltmP*, *ltmQ*, *ltmG*, *ltmM*, *ltmC*, *ltmD* and *ltmB* are predicted to be orthologues of the *P. paxilli paxP*, *paxQ*, *paxG*, *paxM*, *paxC*, *paxD* and *paxB*, respectively. The genes *ltmE*, *ltmJ* and *ltmK* have no known *pax* orthologue and are predicted to be unique to lolitrem biosynthesis. No homologue of the *P. paxilli paxA* gene was identified within the three gene clusters.

The proposed lolitrem B biosynthesis pathway (Fig. 4.1; 4.2) has been based on identification of indole-diterpene compounds identified from endophyte-infected perennial ryegrass (Munday-Finch, et al., 1998; Gatenby, et al., 1999; Parker, Scott, 2004; Mantle, Weedon, 1994; Munday-Finch, et al., 1996b). The similarity between paxilline and lolitrem biosynthesis suggests that the gene products of *ltmG*, *ltmM*, *ltmC*, *ltmP* and *ltmQ* play the same respective catalytic roles as their *pax* orthologues. The predicted catalytic functions of the *ltm* gene products are indicated in Table 4.1. The GGPP synthase, LtmG, is proposed to synthesise the GGPP moiety and is considered the first step of lolitrem biosynthesis (Table 4.1). Parker and Scott (2004) have proposed that the FAD-dependent monooxygenase LtmM catalyses the epoxidation of GGPP and this is then cyclised by the prenyl transferase LtmC followed by the addition of indole-3-glycerol phosphate. On the other hand, Fueki et al. (2004) have used ^{2}H-

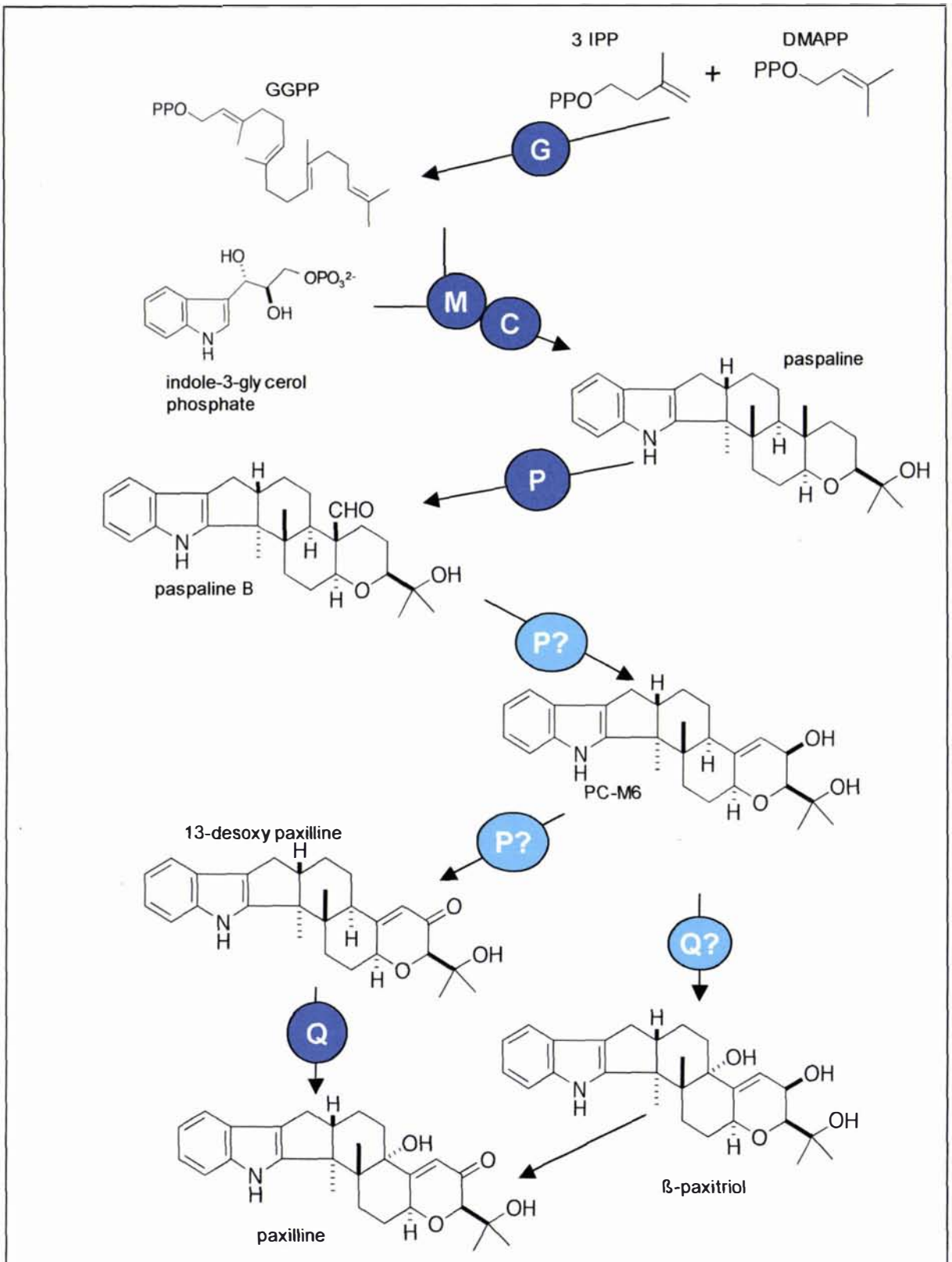


Figure 4.1 Early steps in lolitrem B biosynthesis

The early steps in lolitrem B biosynthesis. The gene names have been abbreviated to a single letter. The proposed enzymatic functions of each gene based on similarity to the *pax* orthologue and deletion analysis is indicated by a dark blue circle. The enzymatic steps indicated in the light blue circles are putative. Paspaline and paxilline are predicted to feed into the lolitrem B metabolic grid (Fig. 4.2). The enzymatic functions of the gene products are found in Table 4.1.

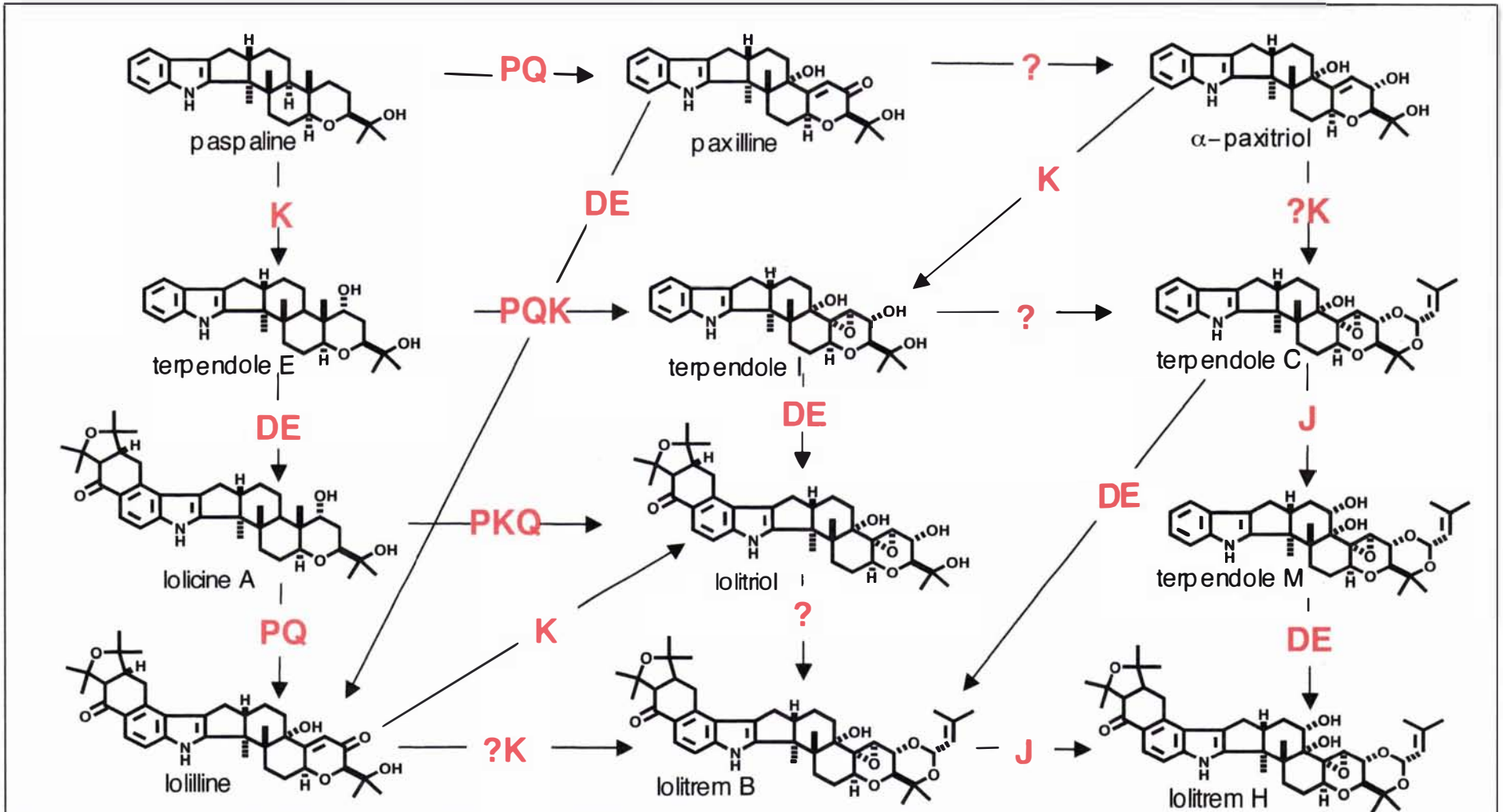


Figure 4.2 The proposed metabolic grid to lolitrem B

The proposed metabolic grid to lolitrem B has been modified from Parker and Scott (2004). The Ltm biosynthetic enzymes have been abbreviated to a single letter. Their proposed catalytic function is described in Table 4.1. The order of the catalytic steps that are predicted to contain two enzymes may result in the synthesis of additional compounds that are yet to be identified. The compounds that have been isolated from endophyte-infected *L. perenne* are paspaline, paxilline, lolicine A, lolitriol, terpendole M, lolilline, lolitrem B and lolitrem H.

Table 4.1 Description of the predicted catalytic activities of the Ltm proteins

Protein	Classification	Predicted Function
LtmE	prenyl transferase/DMAT synthase	Prenylation of C20/ forms AB rings
LtmJ	P450 monooxygenase	Hydroxylation of C14
LtmP	P450 monooxygenase	Demethylation of the C30
LtmQ	P450 monooxygenase	Hydroxylation of the C13
LtmD	DMAT synthase	Prenylation of C21
LtmC	Prenyl transferase	joins indole-3-glycerol phosphate to GGPP
LtmB	Unknown function	
LtmG	GGPP synthase	generates GGPP
LtmM	FAD dependent monooxygenase	cyclisation of GGPP (forms the epoxides
LtmK	P450 monooxygenase	forms epoxide across C11/C12

labelled geranylgeranylindole and shown conversion of this compound to emindole via epoxidation and cationic cyclisation, suggesting that GGPP and indole-3-glycerol phosphate are joined prior to epoxidation. The actual order of the initial biosynthetic steps to paspaline, the first proposed stable intermediate, remains uncertain as stable intermediates have not been isolated from deletion mutations of *paxG*, *paxM*, *paxC* or *ltmM* (Young, et al., 2001; McMillan, et al., 2003; Scott, McMillan, Astin, Saikia, Young, Bryant and Parker unpublished results). Deletion of *ltmM* showed that it was essential for lolitrem biosynthesis as no lolitrems were detected in extracts of perennial ryegrass infected with the *ltmM* deletion transformants (Section 3.6). The *ltmM* deletion mutation was complemented with a native *N. lolii* Lp19 *ltmM* and with *paxM* but only when *paxM* was present in a high copy number. When *ltmC* and *ltmM* were placed under the control of a *paxM* promoter they complemented the *P. paxilli paxC* and *paxM* deletion mutants, confirming that these two genes are functional orthologues of *paxC* and *paxM*, respectively (Section 3.6 and 3.9.4).

The biosynthesis of lolitrems is considerably more complex than that of paxilline, given the diversity of compounds identified in endophyte-infected perennial ryegrass seed and leaf tissue (Gatenby, et al., 1999; Munday-Finch, et al., 1996b; Munday-Finch, et al., 1997). To accommodate this huge structural diversity, Munday-Finch have proposed that lolitrem biosynthesis occurs by way of a metabolic grid with closely related compounds serving as substrates for the same enzyme (Gatenby, et al., 1999; Munday-Finch, et al., 1996b; Munday-Finch, et al., 1997). Therefore, some of the lolitrem biosynthesis enzymes are likely to have broad specificities and catalyse multiple steps via major and minor pathway routes. The *F. fujikuroi* gibberellin biosynthesis pathway requires at least four P450 monooxygenases where P450-1 and P450-4 catalyse multiple steps of the major biosynthetic pathway and also produce side products via a minor pathway route (Rojas, et al., 2001; Tudzynski, et al., 2001; Tudzynski, et al., 2002; Rojas, et al., 2004).

Based on the biosynthetic scheme suggested by Munday-Finch, catalytic roles for the *ltm* gene products can now be proposed (Fig. 4.1; 4.2; Table 4.1). The *P. paxilli* deletion mutants, *paxP* and *paxQ,* accumulate paspaline and 13-desoxypaxilline, respectively. PaxP is proposed to catalyse the demethylation of position C-12 and possibly the oxidation of position C-10 (McMillan, et al., 2003) resulting in either PC-M6 or 13-desoxypaxilline (Fig. 1.3). PaxQ is proposed to catalyse the hydroxylation of

C-13 using either 13-desoxypaxilline or PC-M6 as a substrate (Fig. 1.3). Similar modifications of paspaline are required for lolitrem B biosynthesis. Therefore, LtmP and LtmQ are proposed to act as functional orthologues of PaxP and PaxQ, respectively. To generate lolitrem B, LtmP can catalyse the C-12 demethylation of paspaline forcing the pathway via paxilline or it can demethylate lolicine A shunting the pathway via lolilline (Fig. 1.3, 4.2). The catalytic step from lolicine A to lolilline requires C-12 demethylation and C-13 hydroxylation, thereby requiring both LtmP and LtmQ. Lolitrem B is hydroxylated at the C-13 position, which can be catalysed by LtmQ from at least three different substrates, 13-desoxypaxilline, terpendole E or lolicine A (Fig. 4.2). Based on this scheme it is proposed that deleting either *ltmP* or *ltmQ* would result in mutants unable to produce lolitrem B. The phenotypes of either *ltmP* or *ltmQ* or *ltmP*/*ltmQ* double deletion mutations are predicted to accumulate lolicine A as the metabolic grid would be blocked via paxilline.

The *ltm* clusters 1 and 3 contain *ltmK* and *ltmJ* encoding P450 monooxygenases that are unique to lolitrem biosynthesis. These enzymes are predicted to catalyse steps that have not been identified in paxilline biosynthesis. Additional enzymes are required not only to form the epoxide between C-11 and C-12, identified in lolitriol, lolitrem B and H, and terpendole C, I, and M, but also for the C-14 hydroxylation of terpendole M and lolitrem H. The *ltmK* gene was present in all *N. lolii* strains tested, whereas *ltmJ* was absent in most *N. lolii* strains that are known to be lolitrem B negative, including AR1, Lp14 and *N.* LpTG-2 strain Lp1. *N. lolii* Lp14 produces janthitrems, a structurally similar compound to lolitrem B (Tapper and Lane unpublished) (Fig. 4.3), but indole-diterpenes have not been detected in *N. lolii* AR1 and *N.* spp. Lp1 strains (Latch, et al., 2000). LtmK, an enzyme that is present in both *N. lolii* Lp19 and *N. lolii* Lp14, possibly forms the epoxide moiety of 11, 12-epoxy-janthitrem G, identified in Lp14-infected perennial ryegrass. LtmJ is, therefore, predicted to hydroxylate position C-14, a chemical modification not identified in janthitrem extracts from Lp14. The C-11 hydroxylation that converts paspaline to terpendole E is predicted to be catalysed by a P450 monooxygenase. LtmK is a likely candidate for this role as epoxide formation could not occur prior to C-12 demethylation and, therefore, hydroxylation would occur at the C-11 position. It is now proposed that a deletion mutant of *ltmK* would accumulate the compounds α-paxitriol and lolilline, whereas a deletion of *ltmJ* could still produce lolitrem B. A deletion mutant of both *ltmJ* and *ltmK* would accumulate α-paxitriol and lolilline, the same compounds as that of an *ltmK* mutation. The

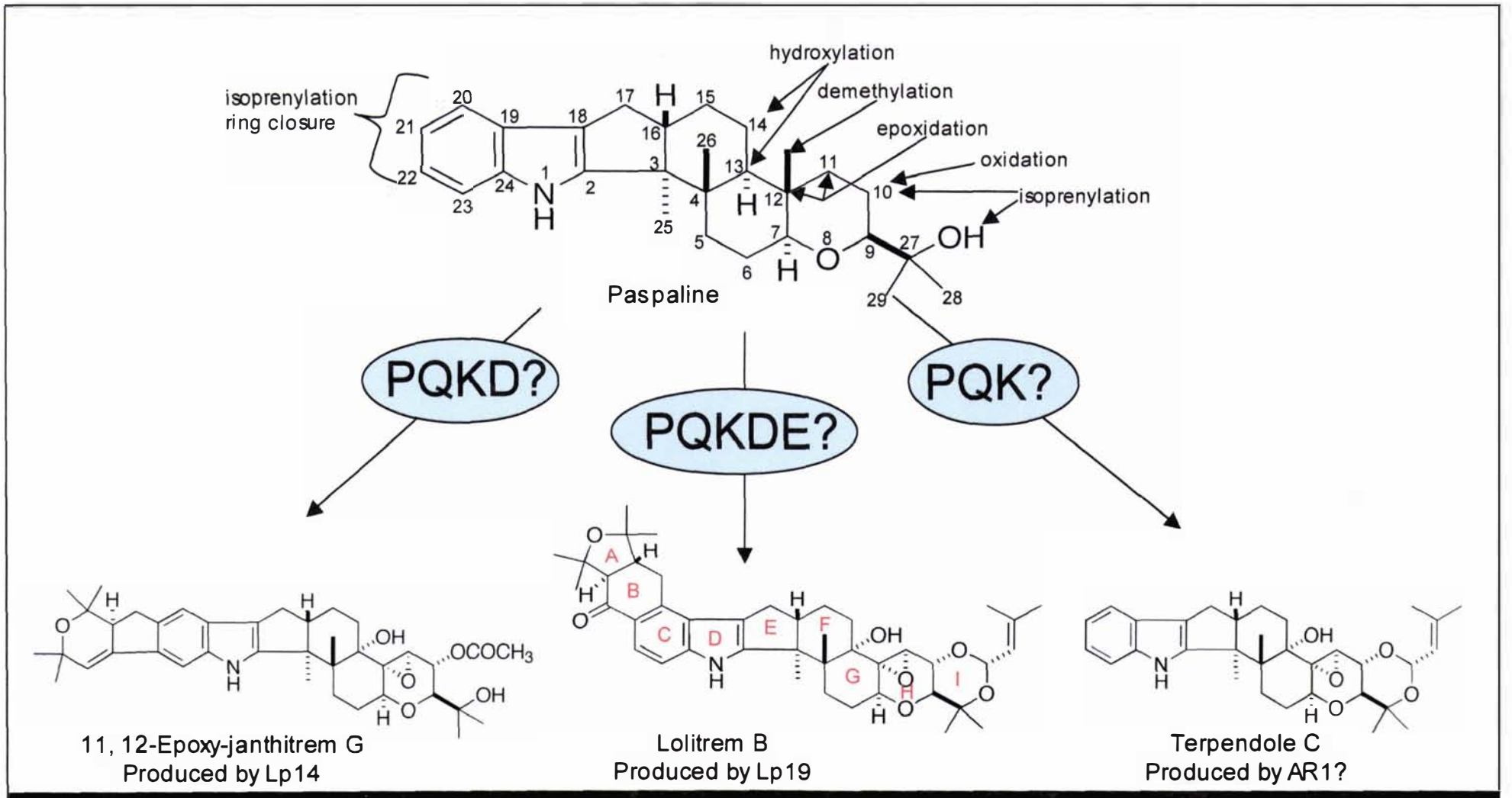


Figure 4.3 Modifications of paspaline and the predicted indole-diterpene phenotypes Lp14, Lp19 and AR1

The modifications to paspaline required to synthesise lolitrem B, terpendole C and 11,12-epoxy-janthitrem G. The indole-diterpene phenotype of AR1 has been predicted based on the presence of *ltm* clusters 1 and 2 assuming that the genes are complete and encode functional proteins. The ring structure of lolitrem B is indicated by the red letters within each ring. Additional genes (represented as a ?) may be required to form the structurally different AB ring found in the janthitrems, and for the formation of the I ring of lolitrem B and equivalent ring structure found in terpendole C.

characterisation of deletion mutants of each of the four P450 monooxygenases would be invaluable for confirmation of the biosynthesis and proposed metabolic grid of lolitrem B.

LtmG, LtmC, LtmD and LtmE are prenyl transferases. The catalytic activities of LtmG, a GGPP synthase, and LtmC, a putative cyclase, are discussed above. LtmD is a putative DMAT-like synthase whereas LtmE contains both a putative DMAT-like synthase and a cyclase domain. The cyclase domain is most similar to LtmC and PaxC whereas the DMAT synthase domain is more similar to LtmD and PaxD than to DmaW (Wang, et al., 2004; Tsai, et al., 1995; Tudzynski, et al., 1999) involved in ergot alkaloid production. Therefore, the *ltm* clusters contain two genes, *ltmC* and *ltmE*, encoding cyclases, and two genes, *ltmD* and *ltmE*, encoding DMAT-like synthases. The functional role of DmaW in ergot alkaloid biosynthesis is the prenylation of the C-4 position of the tryptophan indole ring thereby forming DMAT (Wang, et al., 2004; Tsai, et al., 1995; Gebler, Poulter, 1992; Gebler, et al., 1992). The function of the *P. paxilli* PaxD is yet to be elucidated, but PaxD is predicted to isoprenylate paxilline. Lolitrem B has additional isoprenylation that results in the formation of the A, B and I rings (Fig. 4.3). The isolation of lolicine A, lolitriol, terpendole C and terpendole M from *N. lolii*-infected ryegrass suggests that the A and B rings of lolitrem B can form independently of ring I formation (Munday-Finch, et al., 1998; Gatenby, et al., 1999), requiring at least two enzymes for these prenylations. Lolitrem B, lolitrem H, lolicine A, lolilline and lolitriol (Fig. 4.2) are di-isoprenylated at the C-20 and C-21 positions of the indole ring. The isoprenylation of the C-20 position of lolitrem B is equivalent to that of the tryptophan C-4 position mediated by DmaW in ergot alkaloid production. Therefore, at least one, if not two DMAT synthases, such as LtmD and LtmE, are likely to catalyse such a reaction. The closure of rings A and B formed by the di-isoprenylation of ring C is predicted to be catalysed by a cyclase such as that of the cyclase domain of LtmE or an as yet unidentified P450 monooxygenase/s. A deletion mutant of *ltmE* and/or *ltmD* would lead to terpendole I. The production of janthitrems by Lp14, an isolate that lacks the *ltm* cluster 3, (Fig. 4.3) suggests that LtmD is likely to prenylate the C-21 position. The enzyme required for the prenylation of ring I has yet to be identified.

The formation of lolitrem B can proceed via an intermediate that has been isoprenylated to form the I ring and then subsequently di-isoprenylated to form the A and B rings, or via intermediates that form the A and B rings first and then form the I ring. The P450

monooxygenases would use substrates from both isoprenylation routes, suggesting that the lolitrem B biosynthetic pathway is a very complex grid of reactions that may involve both major and minor routes. Metabolites produced from biosynthetic pathways that involve metabolic grids show greater structural diversity because of these multiple pathway routes. In particular, if enzymes such as P450 monooxygenases are able to catalyse multiple steps and intermediates of these steps are subsequently available for other pathway enzymes, then the structural diversity of the pathway would be magnified with no additional enzymes required. If this hypothesis is correct, additional pathway intermediates are yet to be identified in lolitrem biosynthesis.

The three *ltm* biosynthesis gene clusters contain 10 genes with putative roles in indole-diterpene biosynthesis. However, the presence of additional genes cannot be discounted, especially as the DNA sequence was not extended beyond *ltmE* in cluster 3. Enzymatic steps that are unaccounted for are those required for the conversion of paxilline to α-paxitriol, the formation of ring I in lolitrem B and the prenylation of C-22 in janthitrems. Additional pathway genes could be unlinked, as characterisation of trichothecene production has identified three genes that are not associated with the major gene cluster (Meek, et al., 2003; Peplow, et al., 2003a; Kimura, et al., 1998b; McCormick, et al., 1999; Brown, et al., 2003). Therefore, additional genes may also be recruited from other genomic locations.

4.3 The three *ltm* gene clusters are contained on a 100 kb 'platform'

The three *ltm* gene clusters were shown by Southern and sequence analysis to be linked together on a 100-kb region of the genome (Section 3.9.5). The *ltm* clusters 2 and 3 were demonstrated to be linked by the mapping of overlapping lambda clones that contained a series of direct repeats enclosed in 16 kb AT-rich sequence that separated the two gene clusters. Southern analysis confirmed that cluster 1 is divided from cluster 2 by ~35 kb of sequence predicted to be a 'platform' of retrotransposon elements. Analysis of chromosomal separations showed that the three *ltm* clusters are duplicated on two *N. lolii* Lp19 chromosomes of differing sizes of ~10 Mb and 2.5 Mb. The chromosomal duplication seen in Lp19 is similar to that of *N. lolii* strain Lp5.

However, in Lp5 one copy of *ltm* cluster 3 was translocated from a 2.5 Mb to a 3.5 Mb chromosome.

A transformation of *N. lolii* Lp19 protoplasts with the pCY39 construct, successfully used to generate *E. festucae* Fl1 *ltmM* deletion mutants, did not result in Lp19 *ltmM* deletion mutants (Section 3.6). A total of 111 *N. lolii* Lp19 hygromycin-resistant transformants were screened with the primer combinations used to detect *E. festucae* Fl1 transformants with homologous recombination of the pCY39 plasmid. However, the primer combinations used for screening the transformants would be unable to distinguish a homologous recombination event contained in a duplicated region. Unfortunately, the Lp19-*ltmM* transformants were not screened with primers that would detect the presence of a homologous recombination in the presence of a duplicated target gene. However, these data, along with the Southern hybridisation of genomic DNA and chromosomal separations, support the presence of duplications of the *ltm* genes in *N. lolii* Lp19. Although the quality of the *E. festucae* strain Fl1 chromosomal separations is poor, the *ltm* gene clusters in *E. festucae* do not appear to be duplicated, as *ltmM* deletion mutants were obtained at a frequency of ~5%.

The duplication of the *ltm* genes in *N. lolii* Lp19 and *N. lolii* Lp5 is unusual as duplications of other gene clusters, such as AK-toxin, AM-toxin and HC-toxin, have typically been in tandem or associated on the same chromosome (Panaccione, et al., 1992; Tanaka, et al., 1999; Johnson, et al., 2000; Ahn, Walton, 1996; Ahn, et al., 2002; Tanaka, Tsuge, 2000). Both copies of the duplicated genes involved in the production of HC-toxin are known to be active (Panaccione, et al., 1992). However, in the plant pathogen *Alternaria alternata* only one copy of the AK- and AM-toxin genes is active (Tanaka, et al., 1999; Johnson, et al., 2000). It remains to be determined whether both copies of the *ltm* genes in the Lp19 duplication are active.

4.4 Regulation of the *ltm* genes

The endophyte DNA biomass in endophyte-infected perennial ryegrass was estimated by real-time PCR to be between 0.3 to 1.9 % of the total plant biomass in leaf sheaths (Section 3.6.2). The DNA biomass of *N. lolii* Lp19 and *E. festucae* Fl1 contained in endophyte-infected perennial ryegrass were similar, averaging 0.8% and 1% of the total

DNA biomass respectively. Previous studies of endophyte biomass *in planta*, determined using quantitative PCR, estimated the quantity of DNA of the fungal origin from *Neotyphodium* spp. LpTG-2 strain Lp1 to lie in the range 0.14 - 0.6% (Panaccione, et al., 2001). Although this is a lower estimate than that for the Lp19 and Fl1 samples mentioned above, the result is still consistent with an extremely low endophyte to plant DNA biomass ratio.

The expression profiles of the *ltm* genes showed that these genes were highly expressed *in planta* where the production of lolitrem B is most prevalent (Section 3.9.3). Expression of the *ltm* genes was not detected in mycelial extracts. The expression profiles were confirmed by northern analysis, one step RT-PCR and with cDNA pools from random primed RNA (Section 3.5). As expected, the endophyte *tub2* and *chsV* genes were expressed in samples from mycelia and endophyte-infected perennial ryegrass extracts. ChsV is a class V chitin synthase required for the integrity of fungal cell walls. Therefore, the *chsV* gene is expressed in culture and *in planta* (Horiuchi, et al., 1999; Takeshita, et al., 2002; Amnuaykanjanasin, Epstein, 2003; Madrid, et al., 2003). Expression of the *pks* gene, a pseudogene that appears to be nonfunctional, adjacent to *ltm* cluster 1 was not detected in the samples tested.

Coordinated regulation of genes is a key feature of fungal gene clusters (Section 1.8). The genes of the aflatoxin, trichothecene and HC-toxin clusters contain common regulatory motifs within the promoter regions of each gene, which are identified by regulatory proteins that are also contained within the cluster (Hohn, et al., 1999; Pedley, Walton, 2001; Fernandes, et al., 1998). At present there is no evidence that either the *ltm* or *pax* gene clusters contain regulatory genes. However, the *ltm* genes show a similar pattern of regulation and, therefore, are likely to have common motifs within the promoter regions of each gene.

Regulatory motifs within promoter regions can be difficult to identify. Whole-genome comparisons have been used to identify regulatory elements within yeast (Kellis, et al., 2003). However, there was no sequence diversity for this approach to be useful with the DNA sequence of *ltm* cluster 1 from *N. lolii* Lp19 and *E. festucae* Fl1. An alignment of promoter regions from the plant-induced co-regulated *mig2* genes of *Ustilago maydis* revealed regions of similarity that were subsequently shown by reporter analysis to be responsible for inducible activity *in planta* (Farfsing, et al., 2005;

Basse, et al., 2002). Alignment of the *ltm* promoter regions did not identify obvious elements common to all the *ltm* promoters. However, over-represented sequences were detected when searched at the site for Regulatory Sequence Analysis Tools (see Section 2.7). Motifs of particular interest were those that were only present in the *ltm* promoters and not contained in the promoter of the differently regulated *ggs1* gene (Section 3.9.3). Motifs that have been identified within other plant-regulated fungal genes, such as those from the *mig2* genes (Farfsing, et al., 2005), were found in all promoters screened, including *ggs1*. However, the GAL4 motif in the *M. grisea* ACE1 promoter (Bohnert, et al., 2004) was identified only in the *ltmJ* promoter.

The Cys_6Zn_2 binuclear clusters are motifs found in transcriptional regulators that are common to fungal genomes and bind dyad sequences (van Helden, et al., 2000; Todd, Andrianopoulos, 1997). The identification of an over-represented dyad CTAN{10-12}CTA within the *ltm* genes (Section 3.9.3) could indicate regulation by a Cys_6Zn_2 binuclear cluster protein. The expression profiles of the *mig2* genes are similar to those of the *ltm* genes, where the genes are not detected in culture but are highly up-regulated *in planta*. Analysis of the *mig1* and *mig2* promoters identified regions responsible for both negative and inducible regulation (Basse, et al., 2002; Basse, et al., 2000). The identification of putative regulatory motifs now provides a starting point for promoter dissection of the *ltm* genes.

The regulation of the *ltm* genes is significantly different from that of the paxilline biosynthesis genes, as paxilline production is limited to culture growth, while lolitrem B is predominantly produced *in planta*. Evidence for separate regulation of these two pathways comes from the complementation analysis of the *ltmM*, *paxM* and *paxC* deletion mutations (Section 3.6 and 3.9.4). An Lp19 fragment containing the complete *ltmM* gene was able to complement the *ltmM* deletion mutation of CYFI1-M28 in all of the transformants screened. Although complete copies of the *paxM* gene were also used to transform this mutant, only one transformant, which contained the *paxM* gene in an extremely high copy number, was able to produce lolitrem B (Section 3.6). To complement the *P. paxilli paxM* and *paxC* deletion mutations, the *ltmM* and *ltmC* genes were put under the control of the *paxM* promoter. Paxilline production was restored in at least 60% of the transformants indicating that *ltmM* and *ltmC* are orthologues of *paxM* and *paxC* respectively (Section 3.6.5 and 3.9.4). Co-transformation of a geneticin-selectable marker was used in these experiments and it is likely that some

transformants do not contain the target gene, resulting in the lower complementation frequency. Heterologous expression of the indole-diterpene biosynthesis genes was not affected by position-dependent regulation, as found with the aflatoxin biosynthesis gene *nor-1*. Expression of the *nor-1*::GUS fusion was only detected at *nor-1* homologous recombination sites (Chiou, et al., 2002). GUS activity was not detected with integration of *nor-1*::GUS at the *niaD* or *pyrG* locus even under inducing conditions for these genes (Chiou, et al., 2002).

Native *ltmC* (pCY66) and ltmM (pCY40) were not able complement the *P. paxilli paxC* and *paxM* deletion mutations recpectively. The pCY66 plasmid was co-transformed with the geneticin-selectable marker and was shown by PCR to contain the complete *ltmC* gene (Bryant and Scott, unpublished results).

The complementation results of the *ltmM* deletion mutant and the *P. paxilli paxM* and *paxC* deletion mutants support the hypothesis that the *ltm* and *pax* genes are regulated differently. Complementation with the appropriate promoter indicates that these genes are orthologues, accepting the same substrates. The expression profiles of the *ltm* genes suggest that these genes are up-regulated *in planta* thereby requiring input from the plant host.

4.5 Predicting indole diterpene (ID) phenotypes

Alkaloid analysis of *Epichloë* and *Neotyphodium* isolates indicated that the most common alkaloid produced by endophyte-infected ryegrass is peramine, followed by ergovaline, lolines and then the lolitrems (Clay, Schardl, 2002; Siegel, et al., 1990; Christensen, et al., 1993). A question remains as to why such closely related species are unable to produce the complete set of alkaloids. It was hypothesised that the lolitrems are only produced by 10% of the endophyte strains (Siegel, et al., 1990) with *E. festucae* and the asexual derivatives *N. lolii* and LpTG-2 being the only isolates to have tested positive for indole diterpenes (Clay, Schardl, 2002; Schardl, et al., 1994; Bush, et al., 1997; Christensen, et al., 1993). In fact, some members of the closely related *E. festucae* and *N. lolii* groups consist of both lolitrem producers and non-producers. To examine this phenomenon more closely, genomic DNA of 42 *Epichloë* and

Neotyphodium isolates was screened for the presence of the *ltm* genes by PCR and Southern analysis (Section 3.10). This was then used to predict phenotypes based on the presence and absence of each *ltm* gene and correlated to known lolitrem B phenotypes.

Two genes proposed to be involved in the production of the alkaloid loline, *lolA* and *lolC*, were found to be strictly associated with endophyte strains known to produce lolines (Spiering, et al., 2002). However, the correlation of *ltm* genes with lolitrem producers is not as tightly associated as that of *lolA* and *lolC* genes with loline production. The genes from *ltm* clusters 1 and 2 were identified within isolates that are known to be lolitrem B negative. However, *ltmJ* and *ltmE* from *ltm* cluster 3 are the only two *ltm* genes whose presence (or absence) is directly correlated with strains that are able (or unable) to produce lolitrem B.

The *E. festucae* group can synthesise all four alkaloids. However, no single *E. festucae* strain has been identified that produces all four classes of metabolites. Of the seven *E. festucae* isolates screened for the presence of the *ltm* genes, only four, Fl1, Fg1, E189 and Frc7, contained the *pax* orthologues *ltmG*, *ltmM*, *ltmC*, *ltmB*, *ltmP*, *ltmQ* and *ltmD*. Frc5, Fr1 and Frr1 were devoid of any detectable *ltm* genes thereby deeming them indole-diterpene negative. Among the forty-two isolates screened by PCR for *ltm* genes, only three, *E. festucae* strains Fl1 and Fg1 and *N. lolii* Lp19, tested positive for all 10 *ltm* genes. Fl1 and Lp19 are able to produce lolitrem B, but the indole-diterpene phenotype of Fg1 is unknown. Based on these data, if the *ltm* genes contained within Fg1 are complete and functional, this strain would be predicted to produce lolitrem B. The *Neotyphodium* isolates contain the largest number of *ltm* genes but of these only the Lp19 isolate contained all 10 *ltm* genes. It is yet to be confirmed that the detection of an *ltm* gene by PCR correlates with the presence of a complete and functional gene.

Analysis of lolitrem B requires fluorescence detection, while paxilline analysis requires UV detection that has more interference from UV-absorbing compounds that are extracted from endophyte-infected grass. Therefore, lolitrem B is easier to identify in plant extracts than paxilline and other indole-diterpenes that require UV detection. The isolates *N. lolii* AR1, *N.* spp, Lp1 and *N. coenophialum* Tf28 have tested negative for lolitrem B but have not been screened for other indole-diterpenes (Christensen, et al., 1993; Tapper, Latch, 1999). PCR and Southern analysis of these strains have detected

the presence of the core genes, *ltmG*, *ltmM*, *ltmC*, *ltmP* and *ltmQ*, required for paxilline production. Thus, if these genes are functional then all three isolates are predicted to produce paxilline and possibly terpendoles. Predictions such as these can now be tested on available endophyte-infected plant material.

The *N. lolii* strain AR1 has been successfully marketed as an endophyte that deters Argentine stem weevil (*Listronotus bonariensis*) and is free of the mammalian toxin lolitrem B (Tapper, Latch, 1999; Popay, et al., 1999). PCR and Southern analysis of AR1 at the loci of *ltm* cluster 1 and cluster 2 indicate that AR1 has a similar genomic map to *N. lolii* Lp19, except that AR1 lacks *ltm* cluster 3, indicating a breakpoint beyond the *ltmP* gene. The absence of *ltmJ* and *ltmE* in AR1 indicates that this strain should at least produce terpendoles of which terpendoles D-I are non-tremorgenic. However, analysis of each gene for sequence, expression and functionality would be required to confirm that the *ltm* genes of clusters 1 and 2 in this strain are all functional. Terpendole C is a potent tremorgen in mice (Gatenby, et al., 1999) and is therefore unlikely to be produced by AR1 as tremorgenicity has not been indicated with this strain. The *N. coenophialum* strain Tf28 has a similar *ltm* profile as AR1, containing the genes from *ltm* clusters 1 and 2. The Tf28 genomic map is similar to *N. lolii* Lp19 across these two clusters. However, the presence of *ltmP* from cluster 2 is ambiguous as PCR analysis indicates the presence of the gene but a complete copy of this gene is not supported by the Southern analysis. Tall fescue grass planted in the USA is predominantly infected with *N. coenophialum* and does not produce the indole-diterpene lolitrem B. Tf28 is predicted to produce paspaline and possibly other indole-diterpenes if *ltmP* is functional.

The *N. lolii* strain Lp14 produces janthitrems, compounds that are structurally similar to lolitrem B but contain a different structural configuration of the A and B rings (Fig. 4.3). Lp14 has a similar genomic map to AR1 but the alkaloid phenotype indicates that this strain has additional genes that are yet to be identified. Other possibilities are that the genes that are present differ slightly with their specificity or regulation, which results in a different class of indole-diterpene. Therefore, sequence analyses of the *ltm* clusters in this strain will be important to distinguish if any structural differences in the *ltm* genes exist, which would result in a different chemotype.

Echinopogon ovatus, a tufted perennial grass of New Zealand infected with a fungus serologically related to *N. lolii*, tested positive for paxilline analogues by an ELISA-based assay (Miles, et al., 1998), although the indole-diterpene compound was not identified. Analysis of endophyte isolates *N. aotearoae* and *N. australiense* from *E. ovatus* revealed that the *N. australiense* isolate, an interspecific hybrid of *E. festucae* and ETC origin, does not contain *ltm* genes. However, analysis of *N. aotearoae* strain E899 demonstrated that genes for *ltmG*, *ltmM*, *ltmB* and *ltmD* are present, but the absence of *ltmC* would suggest that E899 would be unable to produce paspaline, the first proposed indole diterpene. Further analyses of the *ltm* genes present in endophytes from *E. ovatus* would help to predict and subsequently identify the ELISA cross-reacting compound.

The PCR and Southern analysis for the *ltm* genes showed that, apart from *E. festucae* and the asexual *Neotyphodium* strains that contain an *E. festucae* origin, most *Epichloë* endophytes lacked these genes. Therefore, isolates that do not contain the *ltm* genes are unable to produce indole-diterpenes. The *E. bromicola* isolate E799 tested positive for *ltmB*, *ltmC*, *ltmJ* and *ltmE*. However, to be able to produce an indole-diterpene compound, *ltmG*, and *ltmM* would be essential. The *E. typhina* Poa isolate contained the core genes *ltmG*, *ltmM* and *ltmC* required to produce paspaline, but indole diterpenes have not been detected in this Poa isolate. Further analysis is required to determine whether these PCR fragments were from complete and functional genes. The *Epichloë* isolates that showed no evidence of the *ltm* genes would be unable to make an indole-diterpene compound and therefore the absence of an indole-diterpene phenotype is due to the fungal genotype and not to the plant genotype. The presence of remnant *ltm* genes in some sexual *Epichloë* isolates would suggest that the *ltm* genes were once present in a common *Epichloë* ancestor and have subsequently been lost. These data are consistent with the phylogenomic characterisation of fungal *PKS* genes, which indicates that a common ancestor of the Pezizomycotina contained each class of PKS and these have been lost in some species over time (Kroken, et al., 2003).

Deletions of secondary metabolite pathway genes have been identified in isolates of *A. oryzae* and *A. sojae*, which are closely related to *A. flavus* and *A. parasiticus*. Aflatoxin is a highly carcinogenic compound produced by *A. flavus* and *A. parasiticus*, but is not detected in *A. oryzae* or *A. sojae*. Analysis of *A. oryzae* and *A. sojae* strains showed evidence of aflatoxin biosynthesis genes but expression of these genes was not detected

(Watson, et al., 1999; Kusumoto, et al., 1998). *A. oryzae* isolates were examined for the structure of the aflatoxin biosynthesis gene cluster. Three distinct groups were observed, those that appear to have all the genes, and two groups that contain deletions, which would indicate two independent breakpoints (Kusumoto, et al., 2000). The *Neotyphodium* isolates such as AR1, Lp14, Lp1 and Tf28 contain *ltm* clusters 1 and 2 and have deletions within the low copy repeat between *ltm* clusters 2 and 3. However, the extents of these deletions in the isolates are unmapped. Analysis of the *Neotyphodium* isolates *in planta* is required to determine whether the *ltm* genes that are present in these strains are expressed and active.

To date there has been no single isolate that has produced both indole-diterpenes and lolines and therefore the biosynthesis of these two compounds appears to be mutually exclusive of each other. Comparative analysis of genomic sequence from isolates such as the loline-producing *N. coenophialum* Tf28 and lolitrem-producing *N. lolii* Lp19 will be useful to establish if there is a genetic reason why the capacity to synthesise both lolines and lolitrems in one strain does not occur.

4.6 Features of the *N. lolii* genome

The genomes of *Epichloë* and *Neotyphodium* harbour many repeated sequences that range from low copy numbers (less than 10) to significantly high numbers (greater than 50) as seen with the retrotransposons (Section 3.7 and 3.9). To date, the retrotransposon sequences identified within *Epichloë* and *Neotyphodium* lack open reading frames and contain a sequence bias towards A and T nucleotides. The repetitive sequences that are associated with the *ltm* genes, such as Tahi, Rua and the low copy repeat between *ltm* clusters 2 and 3, have not been identified in *Epichloë* isolates that belong to the ETC group. The sequence analysis of the Tahi and Rua retrotransposons indicated they are degenerate and would be non-functional. The high AT content of the repeated sequences suggested that an RIP-like mechanism maybe functional within *Epichloë* species. RIP is a process that specifically mutates repetitive DNA prior to meiosis, by introducing transitions of G:C to A:T. RIP functions premeiotically (Selker, 1990) and therefore to determine whether RIP is functional within the *Epichloë* sexual species, a sexual cross with a strain containing multiple copies of a selectable marker would be

required. It is unlikely that RIP occurred within the asexual strains and therefore all G:C to A:T transitions have accumulated prior to the loss of the sexual state. This is supported by a sequence identical to Rua-2, a partial retrotransposon sequence that is located adjacent to *ltmG* in the sexual *E. festucae* strain Fl1 and the asexual *N. lolii* strain Lp19. There is also no evidence that RIP has altered the duplicated *ltm* genes in *N. lolii* strain Lp19, suggesting that the duplication of these genes occurred once as the asexual state was formed.

Retrotransposons have been implicated in genomic rearrangements in *Saccharomyces cerevisiae* (Kim, et al., 1998). Repeated sequences within a genome can act as regions of homology, which can serve as sites for ectopic recombination. Ectopic recombination between chromosomes can be responsible for translocations, duplications and deletions (Zolan, 1995). The *Epichloë festucae* and *Neotyphodium* genomes contain both high and low copy number repeats that could act as recombination 'hotspots' for genomic rearrangement. The absence of *ltm* cluster 3 in many *Neotyphodium* isolates that contain clusters 1 and 2 is indication of a recombination 'hot spot'. Therefore, a large volume of dispersed repeats, such as the retrotransposons, can contribute to an increase in the plasticity of the genome (Zolan, 1995). Deletion of sequences is not uncommon to fungal genomes and is known to spontaneously occur at the loci for HC-toxin (Pitkin, et al., 2000), AM-toxin (Johnson, et al., 2001), AK-toxin (Hatta, et al., 2002), *AVR-Pita* (Orbach, et al., 2000) and *PWL2* (Sweigard, et al., 1995). The genes required for the production of AK- and AM-toxin are found on conditionally dispensable chromosomes. CYFl1-M61, generated from the transformation of Fl1 with plasmid pCY39, contains a large deletion associated with the *ltm* genes. The extent of the deletion remains undefined but extends beyond *ltm* cluster 3, which indicates that regions within and between the *ltm* clusters are dispensable and therefore not required for growth of the organism.

The *ltm* gene platform is unique in that retrotransposons and low-copy repeats interrupt a series of genes that were perhaps once contained as a single gene cluster. In support of an ancestral single cluster, the *E. festucae* isolate Fl1 lacks the retrotransposon platform between the *ltm* cluster 1 and the *pks* pseudogene (Section 3.7). The data presented here also show that the *ltm* genes existed before the retrotransposons, as fragments of *ltm* genes are still present in isolates that do not contain the retrotransposons. The *E. bromicola* isolate E799 has no evidence of the

retrotransposons Tahi and Rua, but does contain fragments of *ltmB*, *ltmC*, *ltmJ* and *ltmE* (Section 3.10).

Phylogenetic analysis of the asexual *Neotyphodium* isolates has predominantly found these strains to be interspecific hybrids with at least two ancestral parents (Schardl, et al., 1994; Tsai, et al., 1994; Moon, et al., 2000; Moon, et al., 2004). The most prevalent interspecific hybridisations have been associated with an *E. festucae* ancestor (Moon, et al., 2004). While nearly all the asexual endophytes are interspecific hybrids, *N. lolii* and *N. aotearoae* appear to be the exceptions (Moon, et al., 2004; Schardl, Craven, 2003). The duplication of the *ltm* genes in *N. lolii* isolates Lp19 and Lp5 could be the first evidence that *N. lolii* is an intraspecific hybridisation of two *E. festucae* parents. This hypothesis will be difficult to test, as the resulting diploid genome would result in identical copies of each gene. The genome size of *N. lolii* was previously estimated as between 28-35 Mb as determined by quantitative Southern analysis and electrophoretic karyotypes (Kuldau, et al., 1999). Electrophoretic separation of the *N. lolii* Lp19 and *N. lolii* AR1 chromosomes provides estimates of their genome sizes at approximately 38 Mb and 36 Mb, respectively (Section 3.9.5). This indicates that the genome sizes of these strains are near the upper limits of a haploid genome but consistent with the size of the *E. festucae* isolates used at the same time. However, chromosomal separations are often difficult to analyse as the chromosomal bands do not follow a standard ratio of staining intensity to size. If strains such as *N. lolii* Lp19 have resulted from hybridisation with two *E. festucae* parents, both parents would have contained the *ltm* genes. *N. lolii* isolates such as AR1 and Lp7 that have only one copy of the *ltm* gene clusters could have arisen by hybridisation with two different *E. festucae* isolates, one that contained the *ltm* genes and the other that did not. The *N. coenophialum* isolates have arisen from two separate hybridisations with subsequent reduction of the polyploidy state by deletion, as not every gene is present in triplicate and the genome size of ~61.2 Mb is equivalent to that of a diploid rather than a triploid genome (Moon, et al., 2004; Kuldau, et al., 1999). Comparison of two related yeast species, *S. cerevisiae* and *Kluyveromyces waltii*, has shown that *S. cerevisiae* underwent whole-genome duplication, which was then followed by mutation, deletion and rearrangement to regain a stable ploidy state (Kellis, et al., 2004). The genome size of *N. lolii* is similar to that of the parental ancestor, *E. festucae*. This suggests that if a genomic hybridisation event occurred with two *E. festucae* parents, extensive deletion followed to return the *N. lolii* genome back to a functional size. An alternative hypothesis is that

the *ltm* genes and flanking region in *N. lolii* have been duplicated and translocated to another chromosome. Further sequence analysis of both *E. festucae* and *N. lolii* may resolve this issue.

Karyotype analysis of the eight *Epichloë* and *Neotyphodium* isolates showed chromosome length polymorphisms, although some strains had similar-sized chromosomes. However, the *ggs1* fragment hybridised to different-sized chromosomal DNA bands in each strain tested, indicating that genomic rearrangements had occurred (Section 3.2 and 3.9.5). The *ltm* clusters 1 and 2 are contained on chromosomes of identical sizes in the *N. lolii* isolates, Lp19 and Lp5, but these two gene clusters are contained on a different-sized chromosome in the closely related *N. lolii* AR1 and *E. festucae* Fl1 isolates. Although the isolates have different genomic rearrangements, microsynteny was observed at the *ggs1* locus between *N. lolii* strain Lp19 and the filamentous fungi *M. grisea*, *N. crassa*, *F. graminearum* and *A. nidulans*; a similar observation was made with sequences from a BAC clone of *E. festucae* (Kutil, et al., 2004).

4.7 Summary

This thesis describes the isolation of a 100-kb platform of three gene clusters that together are required for the biosynthesis of the indole-diterpene, lolitrem B. The *ltm* locus contains at least seven orthologues of the *P. paxilli* paxilline biosynthesis gene cluster, *pax*. Functional characterisation of two *ltm* genes, *ltmM* and *ltmC*, showed they are involved in indole-diterpene production and are functional orthologues of *paxM* and *paxC*, respectively. The genes contained within the *ltm* locus are highly expressed *in planta* where the production of lolitrem B is most prevalent. Analysis of DNA biomass from endophyte-infected perennial ryegrass showed that the endophyte only represents approximately 0.3 – 1.9% of the total DNA biomass. Therefore, when taking into consideration endophyte biomass *in planta*, the level of the expressed *ltm* genes is greater than that of the perennial ryegrass ß-tubulin or actin genes. These data suggest that the regulation of the *ltm* genes requires plant-specific factors. The indole diterpenes have the ability to protect the host from grazing insects and herbivores. This is mirrored with the *ltm* genes being highly up-regulated *in planta* in correlation with lolitrem B production.

The capacity to produce lolitrems, as seen in *N. lolii* isolate Lp19 and *E. festucae* isolate Fl1, requires the three *ltm* gene clusters. These gene clusters are nonexistent in many *Epichloë* species, which explains why indole-diterpenes have not been detected in these isolates. The genes contained on *ltm* cluster 3 are rarely identified within endophyte isolates, which correlates with an inability to produce lolitrem B. The deletion of *ltm* cluster 3, common to many *Neotyphodium* isolates, indicates a breakpoint between *ltm* clusters 2 and 3. Genomic profiling of *Neotyphodium* species with the *ltm* genes shows the presence of many *ltm* genes, but indole diterpenes are yet to be identified for most of these isolates. Functional data are now required to show that isolates containing the minimum core *ltm* genes have the capacity to produce indole-diterpenes.

The indole-diterpene biosynthesis pathway is a complex grid of enzymatic steps. It is foreseeable in the future that other indole-diterpenes will be identified from *Neotyphodium* and *Epichloë* isolates. Sequence from the regions flanking the *ltm* platform of Lp19 and other indole-diterpene producing endophytes will ascertain whether additional genes are involved in the chemical modification of the indole-diterpenes. The expression profiles and chemical phenotypes of many endophyte isolates have yet to be determined, which may add further support to the correlation of the *ltm* genes with a chemical phenotype. Comparative genomics and functional analysis of the *ltm*, *pax* and *atm* genes will help elucidate the lolitrem B biochemical pathway.

This is the first description of 10 genes required for the production of the lolitrem indole diterpenes. The identification of these genes has been used to predict the biochemistry for lolitrem production, thereby aiding in the elucidation of the proposed metabolic grid and future identification of additional indole diterpenes compounds. Based on this work most *Epichloë* endophytes can now be classified as 'unable to produce indole diterpenes'. However, *Neotyphodium* isolates that were previously considered lolitrem negative may have the capacity to produce indole diterpenes. Understanding the biochemistry of indole-diterpene production may result in the identification of novel indole-diterpene compounds.

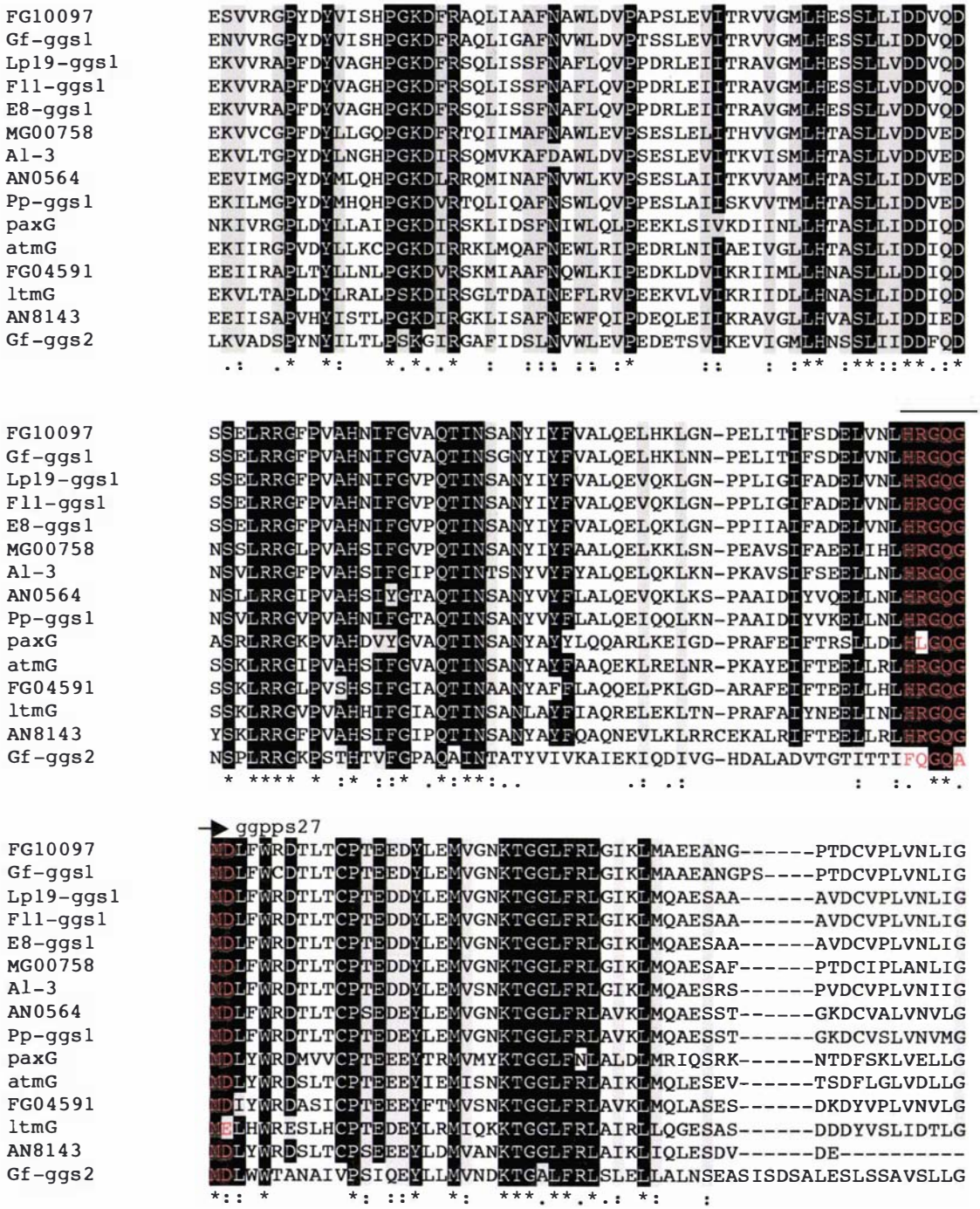
APPENDIX 5.1
MULTIPLE SEQUENCE ALIGNMENTS

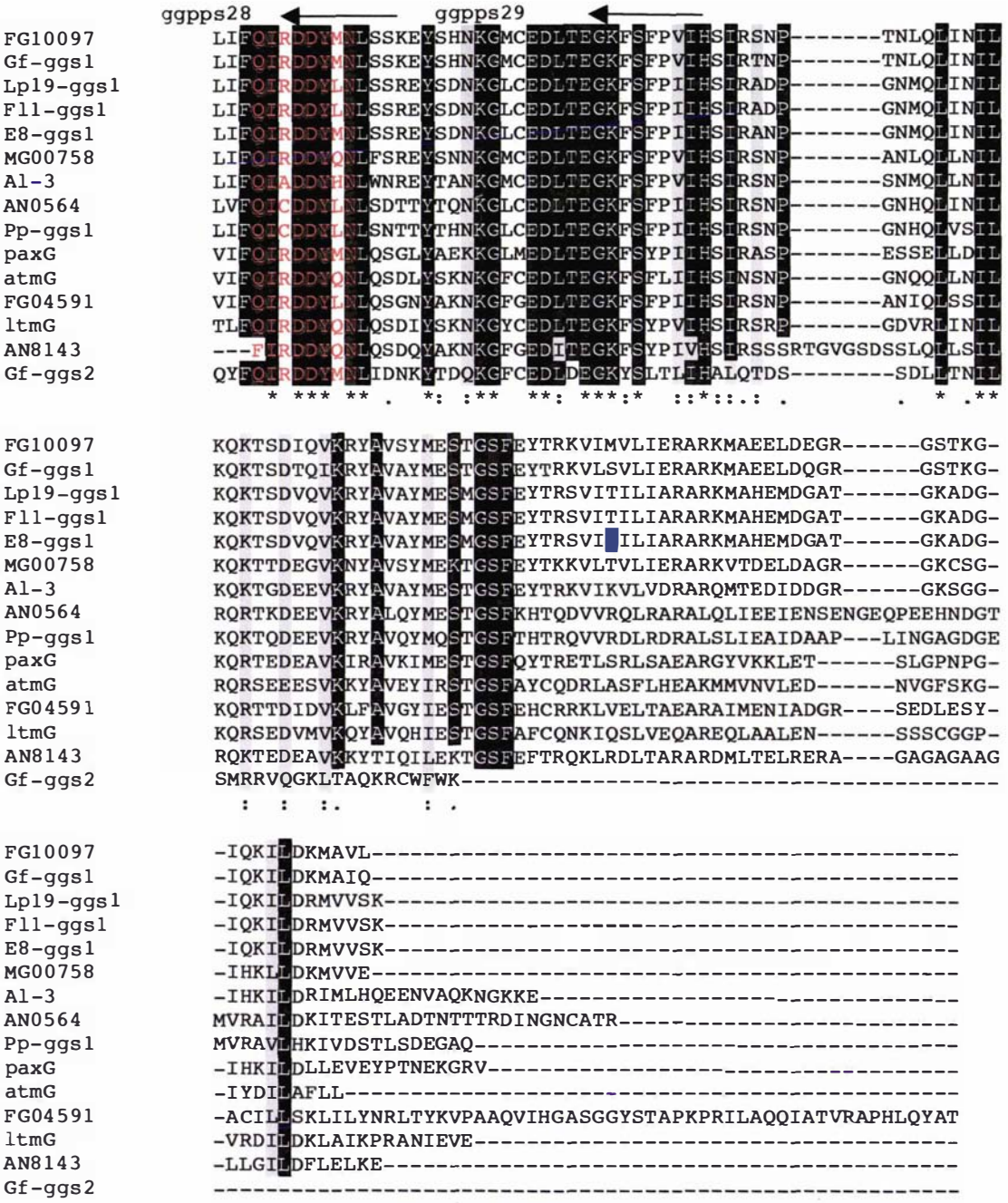
5.1.1 Geranylgeranyl diphosphate synthases
5.1.2 FAD dependent monooxygenases
5.1.3 P450 monooxygenases
5.1.4 Prenyl transferases, paxC-like
5.1.5 Prenyl transferases, dmaW-like
5.1.6 PaxB-like
5.1.7 Geranylgeranyl diphosphate synthases (required for ID production)

Appendix 5.1.1 A clustalx alignment of fungal *ggs* gene

The IPPRXSS motif is highlighted in red. The regions of identity between at least 13/15 the sequences is highlighted in black and regions of similarity are highlighted in grey. The three degenerate primers are indicated by the directional arrows with the letters of the motifs used in red. Amino acid differences found between Lp19, Fl1 and E8 have been highlighted in Blue. The accession numbers for the sequences are as follows: *A. flavus* atmG = AAT65717; *A. nidulans* AN0564 = EAA65430, AN8143 = EAA59165; *E. festucae* Fl1-ggs1 = AAW88517; *E. typhina* E8-ggs1 = AAW88518; *F graminearum* FG10097 = EAA68323, FG04591 = EAA72205; *G. fujikuroi* Gf-ggs1 = CAA65644, Gf-ggs2 = CAA75568; *M. grisea* MG00758 = EAA49100; *N. crassa* Al-3 = CAD70868; *N. lolii* Lp19-ggs1 = AAW88513, ltmG = AAW88510; *P. paxilli* Pp-ggs1 = AAK11525, paxG = AAK11531.

```
FG10097          MIP--------TADPML-SLNPETLPPSTLHMLSLSPKAMHKMSAISNPLVSPNAIPPRT
Gf-ggs1          MIP--------TADPIL-SFNPEALPPSALHMLSLSPKAMEKMSGISNPLVSPNAIPPRT
Lp19-ggs1        MASSSLSTNNISSPPFT-SPNPIPPRRSSTGPRRNSGGAGAGASNPNNSNSNSNNNYYQN
Fl1-ggs1         MASSSLSTNNISSPPFT-SPNPIPPRRSSTGPRRNSGGAGAGASNPNNSNSNSNNKYYQN
E8-ggs1          MASSSLSTNIISSPPFA-SP PIPPRRSSTGPRRNSAGAGVGASNPNNNYQSQSQSQSQS
MG00758          MLSS------ISPNTFLQSPNVIPPRTSSSGIHSNHIAPSKPALRPVPEGDWIAQSLKQQ
Al-3             ----------MAVTSSSPGPAPLSLLSNNDDFIAPFNINTKFPSAIVPPRTSSNQPISVA
AN0564           ----------MTSDSHFHPPHAIPPRISSNRM-------SGASTRDKAALMGNFEKDWLS
Pp-ggs1          ----------MSSS--FQPPSPIPPRTSS----------AGAASGPASPPS---HRHRKS
paxG             ----------MSYILAEALNFVRRGISHLNYWGASHSLSADNYWESNFQGFPRLLSDSS
atmG             ----------MISGVPDRWKVVAS------------SLSSN--LDASYP------TSSS
FG04591          ---------------------------------------MAQS-----VPP------ASSH
ltmG             -------------------------------------MTMAAN-----DFP------FQCQ
AN8143           --------------MVFDTLNAPQPYQEHYPPRWNVHHSSSP------LPNGHSIIPNGH
Gf-ggs2          ----------MAEQQISNLLSMFDASHASQKLEITVQMMDTYHYRETPPDSSSSEGGSLS

FG10097          SSN---------------GVPTSLSPTPTKSVLRPVPEGN--------WLSQKQASAKTH
Gf-ggs1          SST---------------GIPTSLNATPTKPVLRPVPEGD--------WLSQKQLSPRAQ
Lp19-ggs1        QSH--QSHQSHQGYQGYQGYQSFTQPAPLPNILSPVQEGGSLFGAASPTASFQEQTMRTS
Fl1-ggs1         QSH--QSYQSHQSYQGYQGYQSFTQPAPLPNILSPVQEGGSLFGAASPTASFQEQTMRTS
E8-ggs1          QSQGCQSYQSYQGCQSYQGYQSFTQPAPLPNILSPVQEGSCPFGAASPTASFQEQTMRTS
MG00758          QAS----------------IASATAAAPATTRRMRKATSN----------TPPVDADPH
Al-3             IPS---------------NRISSAGLAATQQAQTRKRKAS----------VAQISLPSM
AN0564           KGD-----------------KLQTN-TDLSKRHTRNQSS----------------------
Pp-ggs1          VTD-----------------EPTTNGPSARLKHKRNKSS----------------------
paxG             KAP-----------------STIRTVQVLEDDVDDIAIQY----------------------
atmG             LST-----------------EPIDTRSSSPQGSASTPVDK----------------------
FG04591          GQL-----------------TVG---EVGTSVNSVPYARH----------------------
ltmG             EKK-----------------SYSQPSLVYCNGNIAETYLE----------------------
AN8143           SII-----------------PNGKTAVAVTNGADNQHQHKG----------------------
Gf-ggs2          RYD-----------------ERRVSLPLSHNAASPDIVSQLCFS-----------------

FG10097          ASNPGYGVMTA--------------------PNPP-PDPERYAHEDLEFTA-KRSWNGDK
Gf-ggs1          MSNAGYGVMQA--------------------PNPP-PDPERYAHEDLEFTA-KRSWTDEK
Lp19-ggs1        MRNPTTPIVSAIMHHHQQQQQQQQQQQQQQRPPPPP-PDPRRYEHEDLNYTP-QRSWTDDK
Fl1-ggs1         MRNPTTPIVSAIMHHHQQQQQQQQQQQQQ-RPPPPP-PDPRRYEHEDLNYTP-QRSWTDDK
E8-ggs1          MRNPTTPIASAIMHHHQHQHQHQQ-----RPPPPPLPDPRRYEHDDLNYTP-QRSWTDDK
MG00758          MLNHQT-------------------------QAPTRFDANRYATEDFNFTTTKKTWSEDK
Al-3             LPTSFSPYTMAPQ------------------PPQPPPNPDRFATEDF-FSPSRRTWSEEK
AN0564           ------------------------------------LDGTKYK---------DGKWSQEN
Pp-ggs1          ------------------------------------VDGTKYK---------DGTWSSKN
paxG             ------------------------------------------------------------
atmG             ------------------------------------------------------------
FG04591          ------------------------------------------------------------
ltmG             ------------------------------------------------------------
AN8143           ------------------------------------------------------------KE
Gf-ggs2          ----------------------------------------------TAMSSELNHRWKSQR
```


FG10097
Gf-ggs1
Lp19-ggs1
Fl1-ggs1
E8-ggs1
MG00758
Al-3
AN0564
Pp-ggs1
paxG
atmG
FG04591
ltmG
AN8143
Gf-ggs2
ggpps27


ggpps28
ggpps29
FG10097
Gf-ggs1
Lp19-ggs1
Fl1-ggs1
E8-ggs1
MG00758
Al-3
AN0564
Pp-ggs1
paxG
atmG
FG04591
ltmG
AN8143
Gf-ggs2

Appendix 5.1.2 A ClustalW alignment of fungal paxM-like sequences

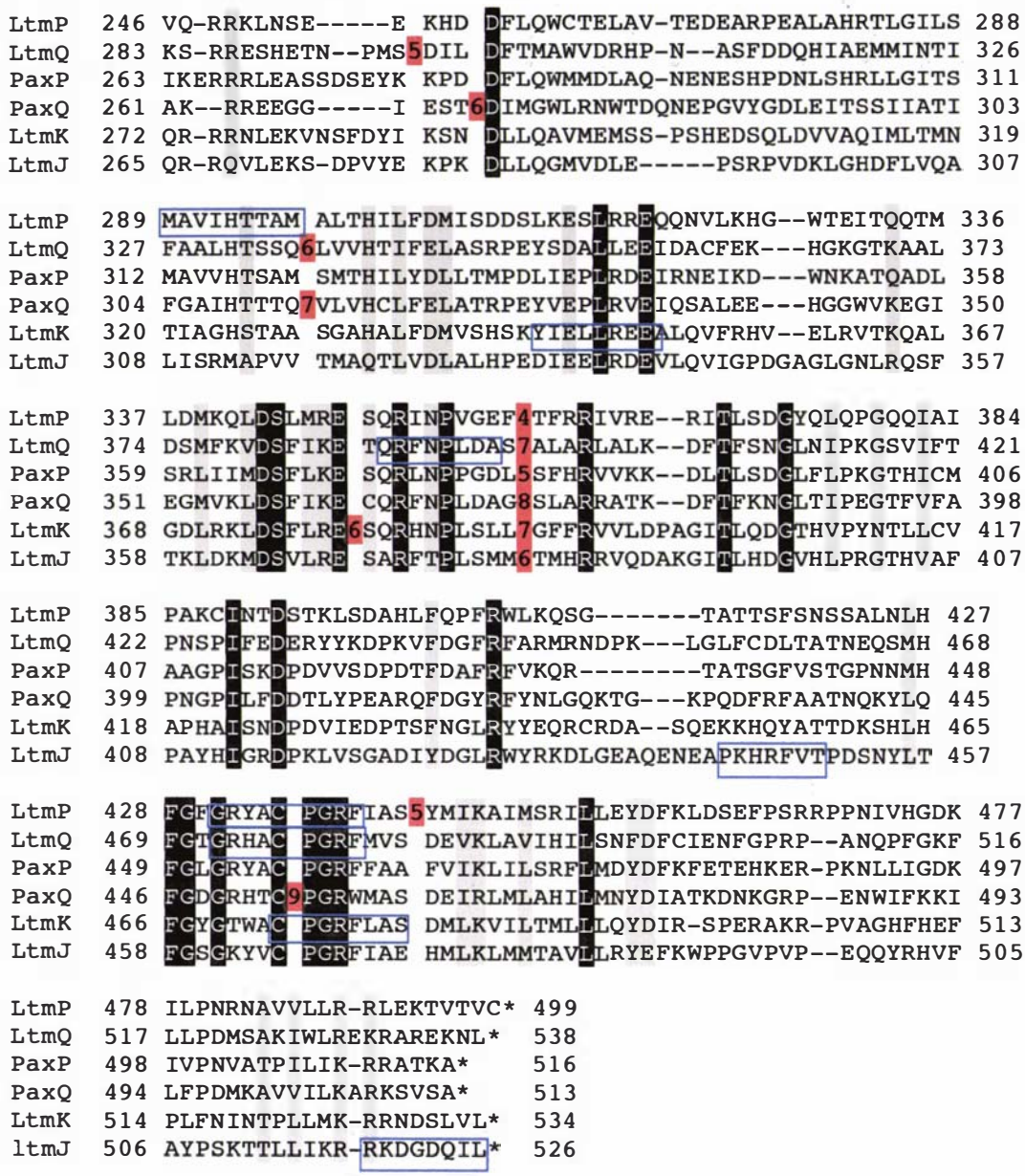
The FAD dependent monooxygenases AtmM from *A. flavus*, PaxM from *P. paxilli* and LtmM from *N. lolii* were aligned using ClustalW. The regions of identity between all three species is highlighted in black and regions of similarity is highlighted in grey. The number and approximate placement of the introns is highlighted in red. The approximate position of the primers, lol158 and lol159, used in Fig. 3.57 is indicated by a blue box.

```
AtmM    MCDKDRFKVIIVGGSVAGLTLAHCLQRAGIDHVVLEKNSDLSPQVGASIGIIPNGGRILD 60
PaxM    -MEKAEFQVIIVGGSIGGLTLAHCLHRAGIKHVVLEKASDPAPQIGASIGILPNGARVLD 59
LtmM    --MTSDFKVIIVGGSVAGLSLAHCLEKIGVSFMVLEKGNQIAPQLGASIGILPNGGRILD 58

AtmM    QLGLFDAVERMTYPLSIATITYPDGYSERNNYPKTVDER1FGYPIAFLDRQKFLEILHTSY 120
PaxM    QLQLYDQVEEHIEPLSKATIGLPDGFNFSSSYPKIIDQR1FGFPIAFLDRQKMLEILYKGY 119
LtmM    QLGIFHSIEDEIEPLESAMMRYPDGFSEKSQYPQALHTS1FGYPVAFLERQRFLQILYDKL 118

AtmM    PDPSNIHTNCRVTHIRRHDSHMEVVTSSGQEYTGDLVVGADGVHSVIRSEMWKLADALEP 180
PaxM    PDPSKIRLGQRVTSIESLDDGVLITTTTGHVYRGDLLVGADGVHSIVRREIWKARGIARR 179
LtmM    KSKDCVFTNKRVVSIASGQDKVTAKTSDGAKYLADIVIGADGVHSIVRSEIWRHLKENSQ 178

AtmM    GRVSKREKR2SMKVEYACVFGISLPVPGLKVGDQVNAFHDGLTIITIHGKNGRVFWFVIKK 240
PaxM    VSKIKQDSS2KLTVEFRCIFGISSAMPGLKLGEQVNALFDGLTIVTIHGKDGRIYWFVIQK 239
LtmM    ISVLEAPNA2SIKHDYSCIYGISLNVPQIILGIQLNCLDDGVSIHLFTGKQSKLFWFVIIK 238

AtmM    LDDMYTYPDTVRFSSADAVRTCENIVHFPLVNGATFGHVWENREVTSMTALEENIFNTWY 300
PaxM    LGKKYVYPDSPRYTSHETSIAAEEIRDVKFYENITFGELWDKRETSSMTALEENTFKVWH 299
LtmM    T-PQASFAKVEIDNTHTARCICEGLRTKKVSDTLCFEDVWSRCTIFKMTPLEEGVFKHWN 297

AtmM    ADRIVCIGDSIH3KMTPNIGQGANTAIEDATVLTNLLYDRLSKNGHKKLARQELLQLLREF 360
PaxM    HGRCVLLGDSVH KMTPNVGQGANMAIEDAAALANLLRKMRISSGPYFPTSSQMEFLLQKY 359
LtmM    YGRLACIGDAIR3KMAPNNGQGANMAIEDACSLANILQKKISHGS---IRDQDINSMFQEF 354

AtmM    QSQRFRRVNKIYQDSRFLVRLHARDGIVKSLLARYIVPYMTELPADLASKSIADSPTIGF 420
PaxM    RDLRYERVNTIYQSSRFLVRFQVRDGIIYSLLSRYWAPYAGDLPADMASKTIADGTMCDF 419
LtmM    SMAQRARTESVCAQSEFLVRMHANQGIGRRLLGRYLIPFLYDAPAGLSGFSISGATRIEF 414

AtmM    LPLPSRSGPGWLQWSRKQRRPATPWILVLLVIVVSFGLHSPELVIPTFWSNSLVSKTVE- 479
PaxM    LPTPKRSGGGWEKYSKQGR--SWSYLTQLMIYLFGLTIVYTSLTMMFDLEGALKFYFLQV 477
LtmM    IDLPTRSLRG--AWGKSWRG-SWEFILQSLVYLRPKFRIVYALYLVAAAAFILYCLSSLF 471

AtmM    -
PaxM    -
LtmM    P 472
```

Appendix 5.1.3 ClustalW alignments of P450 monooxygenases involved in indole-diterpene production

P450 monooxygenases from *N. lolii* (LtmJ, LtmK, LtmP and LtmQ) and from *P. paxilli* (PaxP and PaxQ) aligned using ClustalW. Sequence identity is highlighted in black, and region of similarity is highlighted in grey. The number and aproximate placement of the introns is highlighted in red. The approximate positions of the primers used in Fig. 3.57 are indicated by a blue box.

```
LtmP     1                              MLMLHAVPVGICLLLWYVVYGTK   23
LtmQ     1        MKMLTEHFDFPKLN-----FATIVISGATIIGIIFLRYLNYPTK1  39
PaxP     1 MDLSDFHISTPLRYFHEEAS------LLWKLGVFAVLVYFLLPKPTYKTN   44
PaxQ     1             MDFVLSALQ-----RDSWGIAAIILVSIWALHSFHRSRK   34
LtmK     1 MQYGNLTTVLLLRNTLLSLNSSSICHVHWLQVIVALLVLIVCIFLYWRTP   50
LtmJ     1      MAFASLLHHIWNHAVDCAEQLTWWQTIVSFIIFCIMCSWLPGNGE   45

LtmP    24 RKE CIPTIRRWPRLLPQFLDRLSYNDHAARLVKHGYEK1HKNQPFRLLKMD  73
LtmQ    40 VN- -VPVVGIGVRYT-KWLAAIINVRHARQSIREGYAK2YGDFAFQIPTMT  86
PaxP    45 VK- -VPTVKYMGPWMPEILSRIFFNSHAPTVIYKGYEK1FKTSAFKVVKPD  92
PaxQ    35 LQI1PVPYVGKCGILG-PWISALQWESKARELVQEGYEK2HGNFAFKVALLN  83
LtmK    51 TGI NAPFAGYRSPWEPPLLVQMRYVFNAASMIREGYAK1WKDSLFQISRYD 100
LtmJ    46 MR- -APFVGYRWPFEPTFWVRMRFIFQSLGMMTEGYSK1FKDSMFKITTND  93

LtmP    74 -MDLIVIPLQYALELRAVT-------SDKLD PLTASFD D NAGKVTRILLG 115
LtmQ    87 RMEVFICDRQMTREYQNVDDYHLSFRAVMTE3EFQFKWL L PGQAHEARIIP 136
PaxP    93 -GDLVVLSTRYAEELRQMP-------STTLN ALEATFT D2HVGGYTTILTD 134
PaxQ    84 RWEVCICNEDMIREYKNLMDNQFSAIAVTSE3LFQIKWT A PGTEEGAHKIS 133
LtmK   101 -GDILIVPPRYLDDLHNKS-------QEELS AIYGLIR2N FGGSYSGITLL 142
LtmJ    94 -ADWLVLSQRYLDDLQSLP-------AERLS HTDALVT2M WGSSHSPFALL 135

LtmP   116 SE--LHTRAIQ QRLT-----PKLP2QTLPVLLDELNHAFGQVLPAGN---- 154
LtmQ   137 NSVIAKALSWQ RTRANKPSDPFFE SFSAEFMQGFQEEMRRLIQYQNSSVM 186
PaxP   135 SH--LHTETIQ KKLT-----PAIG3RLIPRMISELDHAFEVEFPTCD---- 173
PaxQ   134 IPLLGKALTWQ RNRS----AAQND PYFSEFVEEFLYAWKEEVPVPENG-- 177
LtmK   143 GENDVGIRALQ3TKIT-----PNLA KLCDDIRDEFQYCLDTDFPACR---- 183
LtmJ   136 NKSDLSSRALR3DVVA-----PNYA KDLDSLVDELRYSLEHDIDIQD---- 176

LtmP   155 --------D G3SNAWISVNPYELVLNLATRATARLFVGDLICRNEIFLETT 196
LtmQ   187 SNRSGAVLD P AHGWHAVPCFPLALKVIGRLTTYVLFGKPLCQDATFLNMC 236
PaxP   174 --------D4- --QFASINPYTVFLRLVARVGARIFIGDELCREEKWLQAS 212
PaxQ   178 --------D4- ----YELPCFETGARVVAHLTARSLVGYPLCRNPEIVNLF 214
LtmK   184 --------D4- ---WTSVSVHPLFLKAVERITHRIFVGLPLCRNPQWVQAT 221
LtmJ   177 --------D4- ---WKPIDALELSSKLVLRISQRILIGWPMSRDQELLECA 214

LtmP   197 ASFSRNT FDTISTSRSFGNLFTH YFARWISTAKEAHGQLQYIQN-LLGSE 245
LtmQ   237 CQFGDVI PRDAIILRSWPALARP4LIVKILSAPR----VMGKLRNILIVEI 282
PaxP   213 IDYTKNI FLTIALMRPMPGFLHP IVGRILPSSRSLKDQLSYIQQDLLGPV 262
PaxQ   215 TDYGSAV PTSGFFIAMFPEIMKP5FVANFCSAPR----ISKRLQAILLEEF 260
LtmK   222 SKHAHYA5TMIQIAMRSVPKFIQP LLNFCLPWPWKNAACVREAKNALILEM 271
LtmJ   215 QGYADAA5TVVQFALKLLPRQIRP LVYPLLPQAWATKSWIRRCDKILAKEM 264
```


LtmP 246 VQ-RRKLNSE-----E KHD DFLQWCTELAV-TEDEARPEALAHRTLGILS 288
LtmQ 283 KS-RRESHETN--PMS5DIL DFTMAWVDRHP-N--ASFDDQHIAEMMINTI 326
PaxP 263 IKERRRLEASSDSEYK KPD DFLQWMMDLAQ-NENESHPDNLSHRLLGITS 311
PaxQ 261 AK--RREEGG-----I EST6DIMGWLRNWTDQNEPGVYGDLEITSSIIATI 303
LtmK 272 QR-RRNLEKVNSFDYI KSN DLLQAVMEMSS-PSHEDSQLDVVAQIMLTMN 319
LtmJ 265 QR-RQVLEKS-DPVYE KPK DLLQGMVDLE-----PSRPVDKLGHDFLVQA 307
LtmP 289 MAVIHTTAM ALTHILFDMISDDSLKESLRREQQNVLKHG--WTEITQQTM 336
LtmQ 327 FAALHTSSQ6LVVHTIFELASRPEYSDALLEEIDACFEK---HGKGTKAAL 373
PaxP 312 MAVVHTSAM SMTHILYDLLTMPDLIEPLRDEIRNEIKD---WNKATQADL 358
PaxQ 304 FGAIHTTTQ7VLVHCLFELATRPEYVEPLRVEIQSALEE---HGGWVKEGI 350
LtmK 320 TIAGHSTAA SGAHALFDMVSHSKYIELLREEALQVFRHV--ELRVTKQAL 367
LtmJ 308 LISRMAPVV TMAQTLVDLALHPEDIEELRDEVLQVIGPDGAGLGNLRQSF 357
LtmP 337 LDMKQLDSLMRE SQRINPVGEF4TFRRIVRE--RITLSDGYQLQPGQQIAI 384
LtmQ 374 DSMFKVDSFIKE TQRFNPLDAS7ALARLALK--DFTFSNGLNIPKGSVIFT 421
PaxP 359 SRLIIMDSFLKE SQRLNPPGDL5SFHRVVKK--DLTLSDGLFLPKGTHICM 406
PaxQ 351 EGMVKLDSFIKE CQRFNPLDAG8SLARRATK--DFTFKNGLTIPEGTFVFA 398
LtmK 368 GDLRKLDSFLRE6SQRHNPLSLL7GFFRVVLDPAGITLQDGTHVPYNTLLCV 417
LtmJ 358 TKLDKMDSVLRE SARFTPLSMM6TMHRRVQDAKGITLHDGVHLPRGTHVAF 407
LtmP 385 PAKCINTDSTKLSDAHLFQPFRWLKQSG-------TATTSFSNSSALNLH 427
LtmQ 422 PNSPIFEDERYYKDPKVFDGFRFARMRNDPK---LGLFCDLTATNEQSMH 468
PaxP 407 AAGPISKDPDVVSDPDTFDAFRFVKQR--------TATSGFVSTGPNNMH 448
PaxQ 399 PNGPILFDDTLYPEARQFDGYRFYNLGQKTG---KPQDFRFAATNQKYLQ 445
LtmK 418 APHAISNDPDVIEDPTSFNGLRYYEQRCRDA--SQEKKHQYATTDKSHLH 465
LtmJ 408 PAYHIGRDPKLVSGADIYDGLRWYRKDLGEAQENEAPKHRFVTPDSNYLT 457
LtmP 428 FGFGRYACPGRFIAS5YMIKAIMSRILLEYDFKLDSEFPSRRPPNIVHGDK 477
LtmQ 469 FGTGRHACPGRFMVS DEVKLAVIHILSNFDFCIENFGPRP--ANQPFGKF 516
PaxP 449 FGLGRYACPGRFFAA FVIKLILSRFLMDYDFKFETEHKER-PKNLLIGDK 497
PaxQ 446 FGDGRHTC9PGRWMAS DEIRLMLAHILMNYDIATKDNKGRP--ENWIFKKI 493
LtmK 466 FGYGTWACPGRFLAS DMLKVILTMLLLQYDIR-SPERAKR-PVAGHFHEF 513
LtmJ 458 FGSGKYVCPGRFIAE HMLKLMMTAVLLRYEFKWPPGVPVP--EQQYRHVF 505
LtmP 478 ILPNRNAVVLLR-RLEKTVTVC* 499
LtmQ 517 LLPDMSAKIWLREKRAREKNL* 538
PaxP 498 IVPNVATPILIK-RRATKA* 516
PaxQ 494 LFPDMKAVVILKARKSVSA* 513
LtmK 514 PLFNINTPLLMK-RRNDSLVL* 534
ltmJ 506 AYPSKTTLLIKR-RKDGDQIL* 526

Appendix 5.1.4 ClustalW alignments of prenyl transferase *ltmC* type sequences

ClustalW Alignment of prenyl transferases (ltmC-type). The sequences are as follows, LtmC and LtmE are from *N. lolii*, PaxC is from *P. paxilli* and AtmC is from *A. flavus*. The regions of identity between all four sequences is highlighted black and regions of similarity are highlighted grey. The number and approximate placement of the introns is highlighted red. Placement of the degenerate primers designed to AtmC and PaxC are boxed in red with the primer name and orientation indicated underneath. The regions boxed in blue shows the approximate placement of primers mentioned in Fig. 3.57.

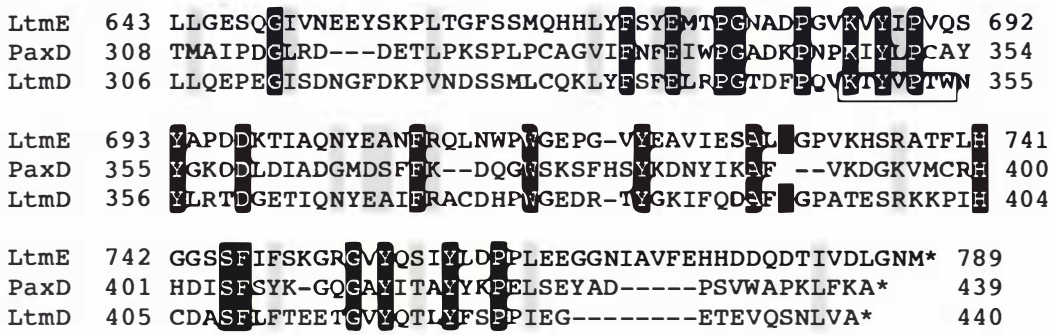
```
LtmC     1 MTSGAWLVARPAAIEIAALLFAFTLGYLVKYTINYQSVVSQAIDHYGYGY  50
LtmE     0                                                     0
PaxC     1                                  MGVAGSGVLYFLFNNVPSPR  20
AtmC     1                MGFFHDFLSRPTTYAILAVLVIPVTALAWDRLP-PL  35

LtmC    51 ERTSHEGIGGSNGK----IPDCPYSYVISLYGHNHFSPLVDFLHPTLKHK  96
LtmE     1            MKP----TTRCPFDYLVSQCGKHHFKTFVQLLSPLLQDE  35
PaxC    21 FWLKKTQLIGTENPEGITGYECPYEYLRKSYGKHHWAAFVDKLSPNLQNE  70
AtmC    36 LPSAKRLLVGKKNPSKITSLECPYSYIRQIYGTHHWAPFVDKLSPSLKTE  85

LtmC    97 YPKKHSLILDIMDAVHLCLIMVDDIICDHSPKRKNHTTAHLLYGSCETANR 146
LtmE    36 DPDRYALILDIMDAVHFSAILIDDIANQSALRRNQPAAHVVFGETETATR  85
PaxC    71 DPAKYRMVLETMDVIHLCLLMVDDIISDGSEYRKGKPAAHKIYGAPETANR 120
AtmC    86 RPAKYHMILEIMDGIHLCLMLVDDIISDGSDYRKGRPAAHHIYGPSETANR 135
                          CY16→                          CY17→

LtmC   147 AYFVLTKVINRAMKEQPVLGIELLRALELILEGQDMSLVWRRDGLRSFES 196
LtmE    86 AYLVLLRVVNRTMRENPVLAGELLNSLEEIHQGQDESLVWRRDGLETFPV 135
PaxC   121 AYYRVTQILAQTATEFPRLSPWLMTDLRDILEGQDMSLVWRRDGVNGFPG 170
AtmC   136 AYYRVTQLLNRTGHEFPELAPWLLQCLEEILEGQDLSLVWRRDGLSAFPV 185

LtmC   197 YGEESLLTYKNMALLKTGTLFVLLGRLLNQGGHQSDDLLGRFG1WYAQLQN 246
LtmE   136 ADDERLAAYVRMSRLKTGSLFVLLGRLLANGGTEFDDLLVRFG1LYAQLQX 185
PaxC   171 TASERTAAYKRMVLLKTGGLFRLLGHLTLENNS-MDEAFSTLG WHSQLQN 219
AtmC   186 QPEERVAAYRQMAYLKTGALFRLVGQLVLENQS-YDDTLSTVA WYSQLQN 234
                          ←CY18

LtmC   247 DCKNIYSEEYAFNKGTV EDLRNRELSFPVVVALNDKHTEPQIRKAFQSQ 296
LtmE   186 DCKNIYSPEYALNKGSV EDLRNGELSYPVVVALIENKAEGIVGEALRTR 235
PaxC   220 DCKNVYSSEYAKMKGVV EDLLNREMTYPIVLALDASGGH-WVEAALKSP 268
AtmC   235 DCKNVYSSDYAKAKGAI EDLRNGELSYPIVVALNVPKGQ-YVVPALEFR 283
        ←CY19

LtmC   297 NQGDIKRALQALESPSVKNTCLKTLQEAGQGLENLVAVWGRKEQMHFTK* 346
LtmE   236 SDGDTEQALRVLESPAVKDACLHALEAASVGLEDLVEAWGRREKMRSDTL 285
PaxC   269 SRRNVGNALKIIQCDYVRDVCMAELARSGAPVKEWLKLVKREEKLDLKA* 318
AtmC   284 SPHNIRQALRVIQSDQVRNICLTEMKKSAVSIQDWLALVGRNEKMDMKSE 333

LtmE   286 DGDDLTRPSTITQHEQDDHVDRAAIDAKSDASGSSNKSLTPPETAPTTDT 335
AtmC   334 K*                                                 335

LtmE   336 LSETAVGDISSVDVDYWTRRCVPIIGSLLKSCRVYSEAERETQLRFLQEH 385

LtmE   386 VLPNLGPRPSSPGSQIQSMATFSGFPLQPSINLSGSGQAKVRYTFEPLDS 435

LtmE   436 LSGTEVDPFALAPAQRVLEKLSTLLGVWPGWIDALIAAYHPTREEVEQLH 485

LtmE   486 PNLHEYLRGVLVRTTGRQDVQVPPMPRMWVCFVALDLEGASQALKVYFDP 535

LtmE   536 KIKEAVTGIPSCKYTCQILRTVDRFGNAKAVDMLEQ2FLAEEHSIGAVELI 585

LtmE   586 AIDCVPEEMQPSARIKVYVHTMSNSFQTVRKYMTMGGRCMDPATLEGLEN 635
```

```
LtmE  636 LHDVWYSLLGESQGIVNEEYSKPLTGFSSMQHHLYFSYEMTPGNADPGVK 685

LtmE  686 VYIPVQSYAPDDKTIAQNYEANFRQLNWPWGEPGVYEAVIESAL[redacted]GPVKHS 735

LtmE  736 RATFLHGGSSFIFSKGRGVYQSIYLDPPLEEGGNIAVFEHHDDQDTIVDL 785

LtmE  786 GNM* 789
```

Appendix 5.1.5 ClustalW alignments of perenyl transferases *ltmD*-like sequences

ClustalW alignment of prenyl transferases (*ltmD*-like). The sequences are as follows, LtmE and LtmD are from *N. lolii*, while PaxD is from *P. paxilli*. The regions of identity between all three sequences is highlighted black and regions of similarity are highlighted grey. The number and approximate placement of the introns is highlighted red. The approximate placement of primers used in Fig. 3.57 are boxed in blue.

```
LtmE     1 MKPTTRCPFDYLVSQCGKHHFKTFVQLLSPLLQDEDPDRYALILDIMDAV  50

LtmE    51 HFSAILIDDIANQSALRRNQPAAHVVFGETETATRAYLVLLRVVNRTMRE 100

LtmE   101 NPVLAGELLNSLEEIHQGQDESLVWRRDGLETFPVADDERLAAYVRMSRL 150

LtmE   151 KTGSLFVLLGRLLANGGTEFDDLLVRFG1LYAQLQXDCKNIYSPEYALNKG 200

LtmE   201 SVAEDLRNGELSYPVVVALIENKAEGIVGEALRTRSDGDTEQALRVLESP 250
PaxD     0                                                  M   1
LtmD     0                                                  M   1

LtmE   251 AVKDACLHALEAASVGLEDLVEAWGRREKMRSDTLDGDDLTRPSTITQHE 300
PaxD     2 QSDELQIPPWESLAEGLG------------------------------  19
LtmD     2 IAKNIELNGLDPATR---------------------------------  16

LtmE   301 QDDHVDRAAIDAKSDASGSSNKSLTPPETAPTTDTLSETAVGDISSVDVD 350
PaxD    20 ------------------------------------------FSNADEE  26
LtmD    17 -----------------------------------------ALDIL  21

LtmE   351 YWTRRCVPIIGSLLKSCRVYSEAERETQLRFLQEHVLPNLGPRPSSPGSQ 400
PaxD    27 YWWTVFGQPLNKLMDWAD-YSTSEKYRVLAFIHRYVIPTCGPRPKP--NG  73
LtmD    22 YWKNHCIKQLESLLCATDSYCTADKAAQLRILSELVLPNLGPRPSN--AT  69

LtmE   401 IQSMATFSGF---PLQPSINLSGSGQAKVRYTFEPLDSLSGTEVDPFALA 447
PaxD    74 DQYWDTFMGFDHTPIQVSINFYNS-KATVRTGNIPISEASGTTEDPINQK 122
LtmD    70 GPSYLTRSGS---PIMLSLNTTSS-KNCVRYCWEILGATGASNDDPLAVQ 115

LtmE   448 PAQRVLEKLSTLLGVWPGWIDALIAAYHPTREEVEQLHPNLHEYLRGVLV 497
PaxD   123 ASLDTIASQRHLV---PGHNLRLFKHFTDAFFIPNEEANILNAELENRTI 169
LtmD   116 VAKDVVASLSATFRLSTKWSETLLSNFAVTPDQARQVINMLPEWIQGFVP 165

LtmE   498 RTTGRQDVQVPPMPRMWVCFVALDLEGASQALKVYFDPKIKEAVTGIPSC 547
PaxD   170 --AM---QAVQ-------CMLSYDFPYRQIQTKVAICPMWKSMQVKRPMG 207
LtmD   166 --EG-MECDFP--KRIPFAMTSFDLNGSNVAMKLYVNPRVKEILTGTPSS 210

LtmE   548 KYTCQILR--TVDRFGNAKAVDMLEQ2FLAEE---HSIGAVELIAIDCVPE 592
PaxD   208 DLMISSIKDLGIDAADYMKSLKVIED FINSEKAVQSGAYAIFFAFDTMLT 257
LtmD   211 DLVWEFLR--NLTPEMKPRAVDLLERIFITDN---SGPSAIELVGIDCVDD 255

LtmE   593 EMQPSARIKVYVHTMSNSFQTVRKYMTMGGRCMDPATLEGLENLHDVWYS 642
PaxD   258 DDYQRTRVKIYFATQSTAFNNMVDIFTLGGRLDGPEMQRATKELRKLWMS 307
LtmD   256 AHLSNARVKLYVHTMSSSFNTVKNYVTLGGAIWDEQTQKGLGILQSIWHL 305
```


LtmE 643 LLGESQGIVNEEYSKPLTGFSSMQHHLYFSYFMTPGNADPGVKVYIPVQS 692
PaxD 308 TMAIPDGLRD---DETLPKSPLPCAGVIFNFEIWPGADKPNPKIYLPCAY 354
LtmD 306 LLQEPEGISDNGFDKPVNDSSMLCQKLYFSFELRPGTDFPQVKIYVPTWN 355
LtmE 693 YAPDDKTIAQNYEANFRQLNWPWGEPG-VYEAVIESALKGPVKHSRATFLH 741
PaxD 355 YGKQDLDIADGMDSFFK--DQGWSKSFHSYKDNYIKAF --VKDGKVMCRH 400
LtmD 356 YLRTDGETIQNYEAIFRACDHPWGEDR-TYGKIFQDSFKGPATESRKKPIH 404
LtmE 742 GGSSFIFSKGRGVYQSIYLDPPLEEGGNIAVFEHHDDQDTIVDLGNM* 789
PaxD 401 HDISFSYK-GQGAYITAYYKPELSEYAD-----PSVWAPKLFKA* 439
LtmD 405 CDASFLFTEETGVYQTLYFSPPIEG--------ETEVQSNLVA* 440

Appendix 5.1.6 ClustalW alignments of PaxB-like sequences

Alignment of PaxB, PaxA (*P.paxilli*) and LtmB (*N. lolii*) using ClustalW. Sequences that are identical between all three strains or between PaxB and LtmB are highlighted black. Identity between PaxA and LtmB is highlighted grey. Approximate intron placements are numbered and highlighted red. The approximate position of the primers used in Fig. 3.56 are boxed in blue.

```
PaxB     1 MDGFD---------VSQAPPEYQAIK-PLADLFVVGM------GV-GWII  33
PaxA     1 MTSITTSVLVLHSLLAANFKYYQSFQNGFIDMLSAMADSNSVSGLPGQLC  50
LtmB     1 MDGFSN--------MEQAPLAYQEVQ-WLAETFVTFM------GL-GWLI  34

PaxB    34 -NYIGM-------------VYISFKHETYGMSIMPLCCNIA----WELVY  65
PaxA    51 REYTGIRPLDTFLTSCTVFFWPTFQGEIPGLSLYGIAFASAMIPMWLIIV 100
LtmB    35 -NYVLM-------------IWHSRRGEPSSMALIPLCNNIA----WELVY  66

PaxB    66 CLVFPSK SPVERGVF-----WMG-LLINFG---VM---YAAITF-SS---  99
PaxA   101 IDVHRRP1TQADRNRFDRLIAFAGPLIQCIGPGLVMPLLLARIHTPSRDSK 150
LtmB    67 TIIYPSP NKVELAAF-----IAG-VTLNFL---IM---TSAARS-AR--- 100

PaxB   100 --REWGHAPLV-ERNISLIF--FVATM---GFLSGHVA---LALEIG-PA 137
PaxA   151 SASQFDYRTFIPSMIIGYILPLLLASLPAPLILSYHNKQQFIAIWQGWPL 200
LtmB   101 --SEWSHSPTM-AKHAGLII--VAGIL---MCFTGHVA---LAMEIG-PA 138

PaxB   138 LA--YSW------GAVICQLLLSVGGLSQLLCRGSTRG----ASYT-L1WA 174
PaxA   201 YSSVLMWAFRRRSGHVHCSRHKGLKHACIFALACSSAGHLVLLSLTWL WS 250
LtmB   139 LA--YSW------GAVICQLALSIGGVCQLLQQHSTGG----TSWK-L1WS 175

PaxB   175 SRFLG----------STCTVG-FAGLRWMY---------WSEAFG-WL 201
PaxA   251 LSYWGYIQSAPWNEPPLASLEAGVLRFLQWDYTLSASATLSWAIAFRHEV 300
LtmB   176 SRFLG----------SCCAVG-FAFLRWRY---------WPEAYG-WL 202

PaxB   202 NSP----LVLWSLVVFLSIDGFYGIC------FWYVDRNEKSLGISGPKK 241
PaxA   301 VQQKSLRISLLSLLRCLIGIVFLGPCSMVALLYWQTCSLQEEAGQKAPAK 350
LtmB   203 ASP----LILWSLATFLVADLT GVC------LLL*                228

PaxB   242 AN*       244
PaxA   351 DKQLDQEI* 359
LtmB   229           228
```

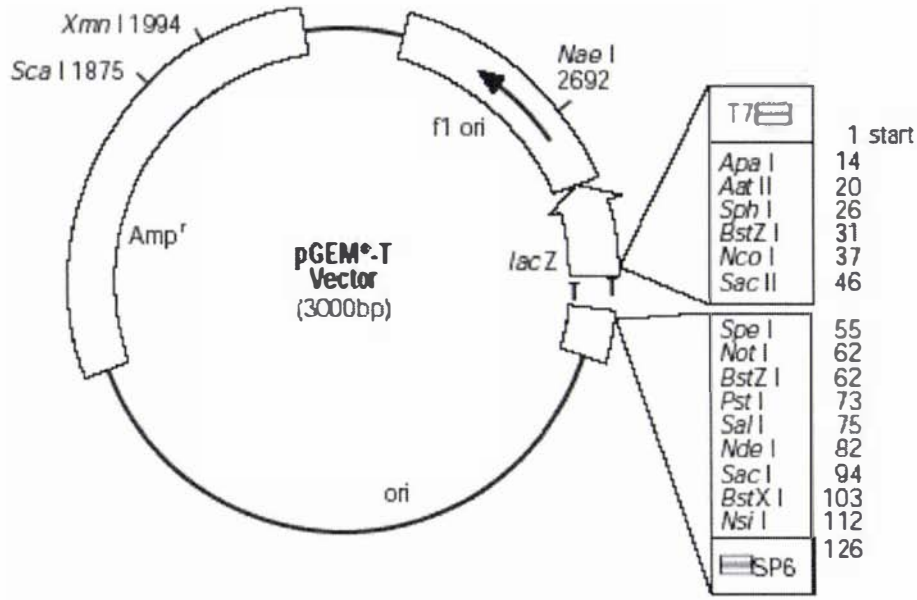
Appendix 5.1.7 ClustalW alignments of Ggs sequences involved in indole-diterpene biosynthesis

Alignment of the three GGPPS involved in indole-diterpene production. The sequences are from the following fungi, LtmG from *N. lolii*, AtmG from *A. flavus* and PaxG from *P. paxilli*. The identical sequences between all three sequences is highlighted in black, similarity is highlighted grey. The introns are numbered and their approximate position is highlighted red. The primers used in Fig. 3.57 are boxed in blue.

```
LtmG     1                                        MTMAANDFPFQCQEKK-   16
AtmG     1                              MISGVPDRWKVVASSLSSNLDASYPTSSSLST-   32
PaxG     1 MSYILAEALNFVRRGISHLNYWGASHSLSADNYWESNFQGFPRLLSDSSK   50

LtmG    17 --SYSQPSLVYCNGNIAETYLEEK1VLTAPLDYLRALPSKDIRSGLTDAI   64
AtmG    33 --EPIDTRSSSPQGSASTPVDKEK1IIRGPVDYLLKCPGKDIRRKLMQAF   80
PaxG    51 APSTIRTVQVLEDDVDDIAIQYNK1IVRGPLDYLLAIPGKDIRSKLIDSF  100

LtmG    65 EFLRVPEEKVLVIKRIIDLLHNASLL2IDDIQDSSKLRRGVPVAHHIFGIA  114
AtmG    81 EWLRIPEDRLNIIAEIVGLLHTASLL2IDDIQDSSKLRRGIPVAHSIFGVA  130
PaxG   101 IWLQLPEEKLSIVKDIINLLHTASLL2IDDIQDASRLRRGKPVAHDVYGVA  150

LtmG   115 QTINSANLAYFIAQRELEKLTNPRAFAIYNEELINLHRGQGMELHWRESL  164
AtmG   131 QTINSANYAYFAAQEKLRELNRPKAYEIFTEELLRLHRGQGMDLYWRDSL  180
PaxG   151 QTINSANYAYYLQQARLKEIGDPRAFEIFTRSLLDLHLGQGMDLYWRDMV  200

LtmG   165 HCPTEDEYLRMIQKKTGGLFRLAIRLLQGESASDD DYVSLIDTLGTLFQI  214
AtmG   181 TCPTEEEYIEMISNKTGGLFRLAIKLMQLESEVTS3DFLGLVDLLGVIFQI  230
PaxG   201 VCPTEEEYTRMVMYKTGGLFNLALDLMRIQSRKNT3DFSKLVELLGVIFQI  250

LtmG   215 RDDYQNLQSDIYSKNKGYCEDLTEGKFSYPVIHSIRSRPGDVRLINILKQ  264
AtmG   231 RDDYQNLQSDLYSKNKGFCEDLTEGKFSFLIIHSINSNPGNQQLLNILRQ  280
PaxG   251 RDDYMNLQSGLYAEKKGLMEDLTEGKFSYPIIHSIRASPESSELLDILKQ  300

LtmG   265 RSEDVMVKQYAVQHIESTGSFAFCQNKIQSLVEQAREQLAALENSSSCGG  314
AtmG   281 RSEEESVKKYAVEYIRSTGSFAYCQDRLASFLHEAKMMVNVLEDNVGFSK  330
PaxG   301 RTEDEAVKIRAVKIMESTGSFQYTRETLSRLSAEARGYVKKLETSLGPNP  350

LtmG   315 PVRDILDKLAIKPRANIEVE*  335
AtmG   331 GIYDILAFLL*            341
PaxG   351 GIHKILDLLEVEYPTNEKGRV* 372
```

APPENDIX 5.2
VECTOR MAPS

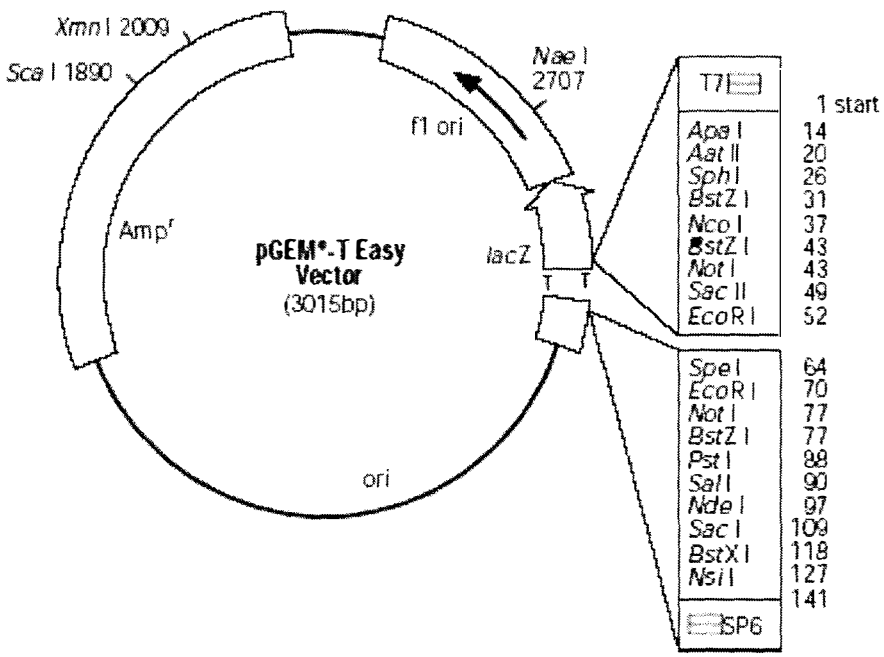
Appendix 5.2.1 pGEMT (Promega)



pGEM®-T Vector sequence reference points:

T7 RNA polymerase transcription initiation site	1
multiple cloning region	10–113
SP6 RNA polymerase promoter (–17 to +3)	124–143
SP6 RNA polymerase transcription initiation site	126
pUC/M13 Reverse Sequencing Primer binding site	161–177
lacZ start codon	165
lac operator	185–201
β-lactamase coding region	1322–2182
phage f1 region	2365–2820
lac operon sequences	2821–2981, 151–380
pUC/M13 Forward Sequencing Primer binding site	2941–2957
T7 RNA polymerase promoter (–17 to +3)	2984–3

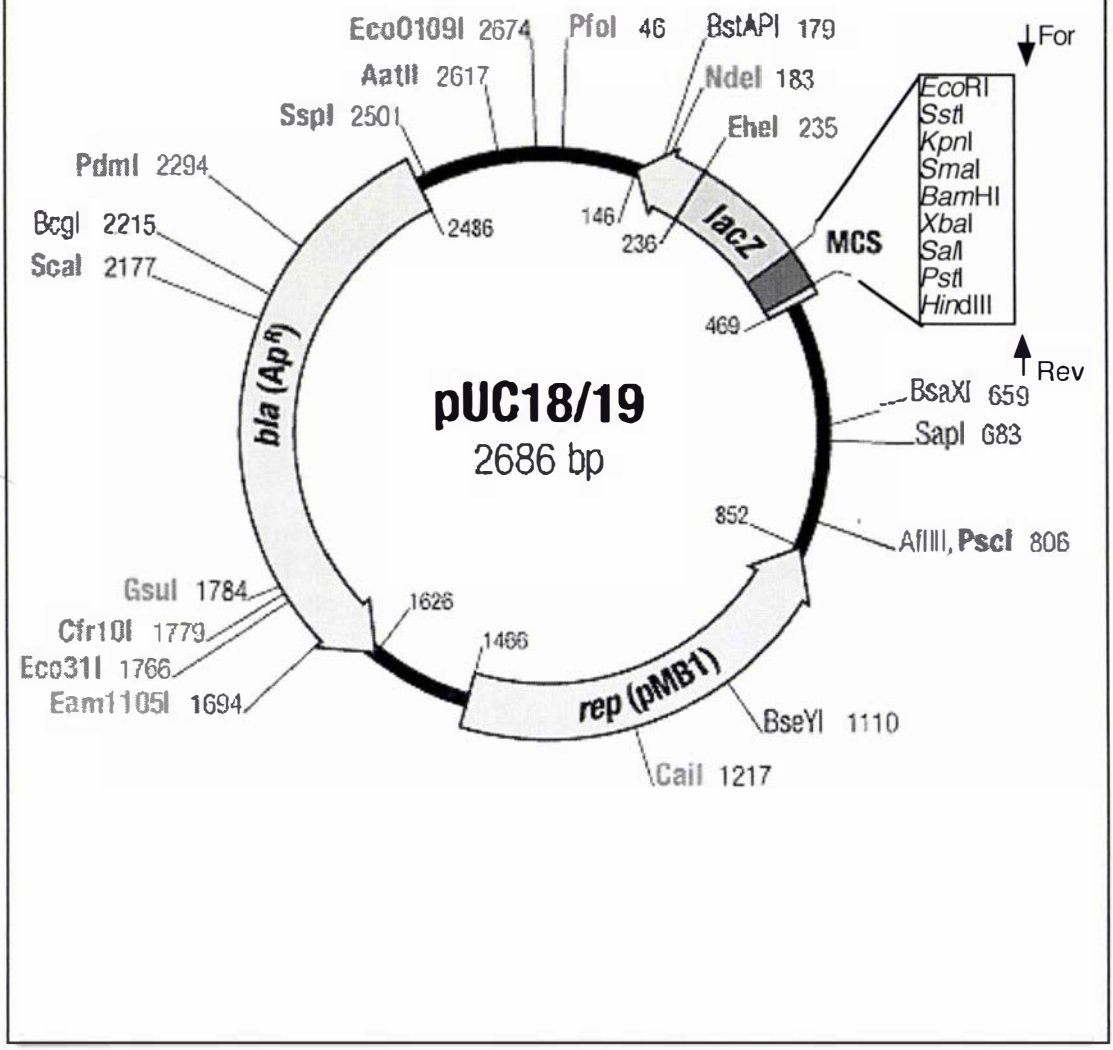
Appendix 5.2.2 pGEM-T Easy (Promega)



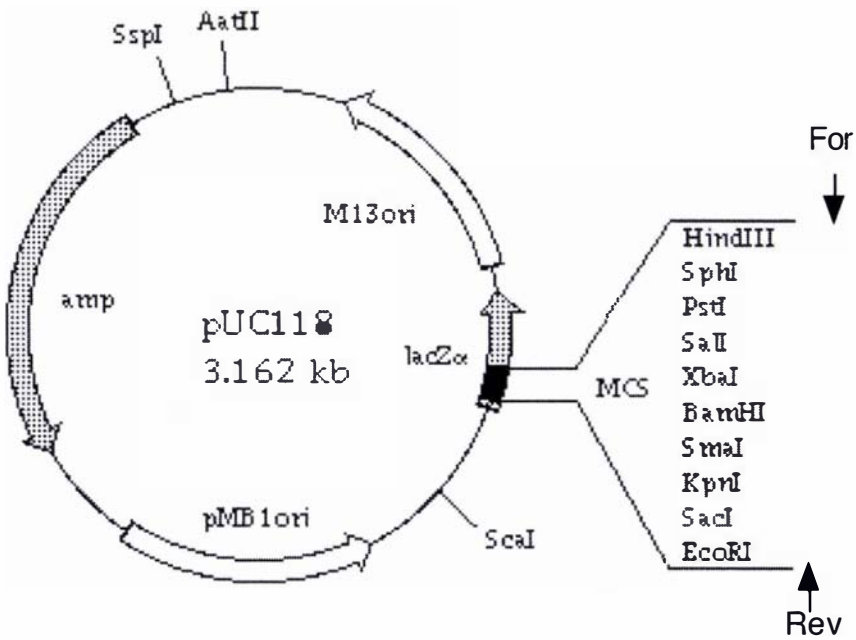
pGEM®-T Easy Vector sequence reference points:

T7 RNA polymerase transcription initiation site	1
multiple cloning region	10–128
SP6 RNA polymerase promoter (–17 to +3)	139–158
SP6 RNA polymerase transcription initiation site	141
pUC/M13 Reverse Sequencing Primer binding site	176–197
lacZ start codon	180
lac operator	200–216
β-lactamase coding region	1337–2197
phage f1 region	2380–2835
lac operon sequences	2836–2996, 166–395
pUC/M13 Forward Sequencing Primer binding site	2949–2972
T7 RNA polymerase promoter (–17 to +3)	2999–3

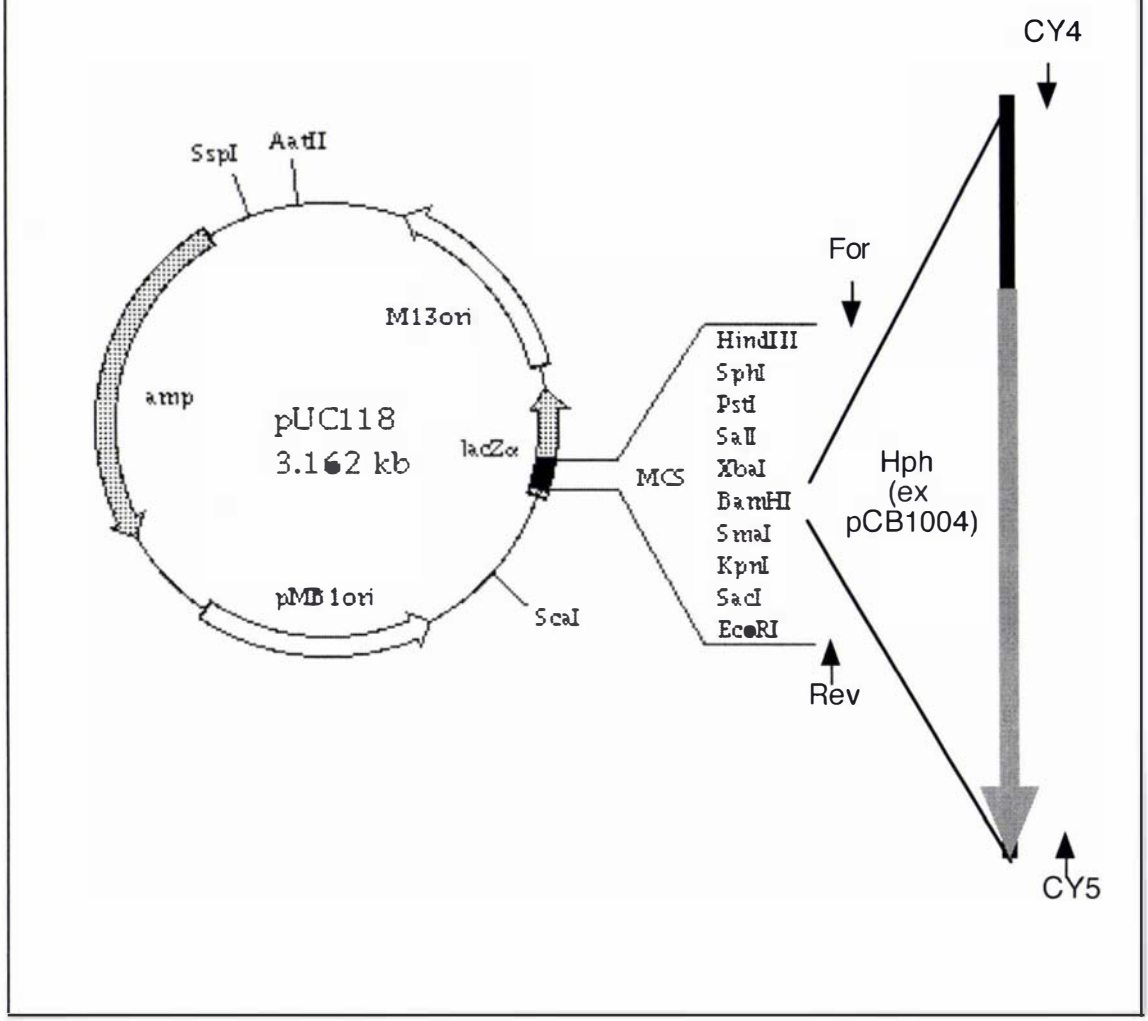
Appendix 5.2.3 pUC19



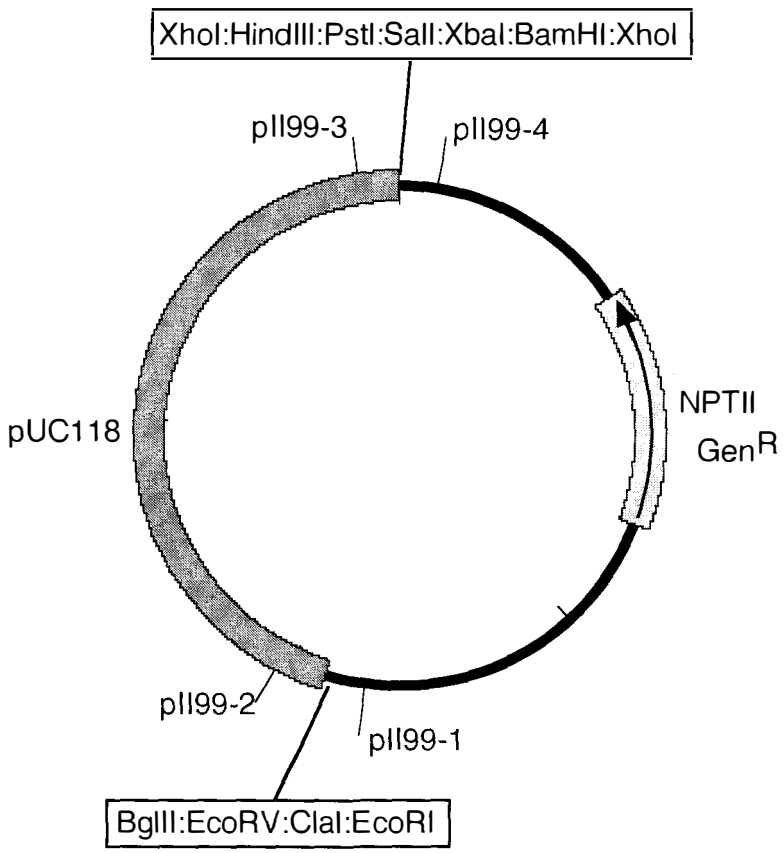
Appendix 5.2.4 pUC118



Appendix 5.2.5 pPN1688



Appendix 5.2.6 pII99



APPENDIX 5.3
SEQUENCE DATA

All sequence data is supplied on the attached CD in sequencher format.

5.3.1 'ggs1' *ggs1* sequences, containing *Nlggs1*, *Efggs1* and *Etggs1*
5.3.2 'CAG' The CAG repeat contained within *ggs1*
5.3.3 'Cluster 1' The *N. lolii* and *E. festucae ltm* cluster 1
5.3.4 'Retro' Sequences of lambda clones isolated with the Rua probe
5.3.5 'Cluster 2 and 3' The *N. lolii ltm* cluster 2 and 3
5.3.6 'PRG tub' Perennial ryegrass tubulin genes
5.3.7 'Cluster 1' MacVector file
5.3.8 'Cluster 2 and 3' MacVector file

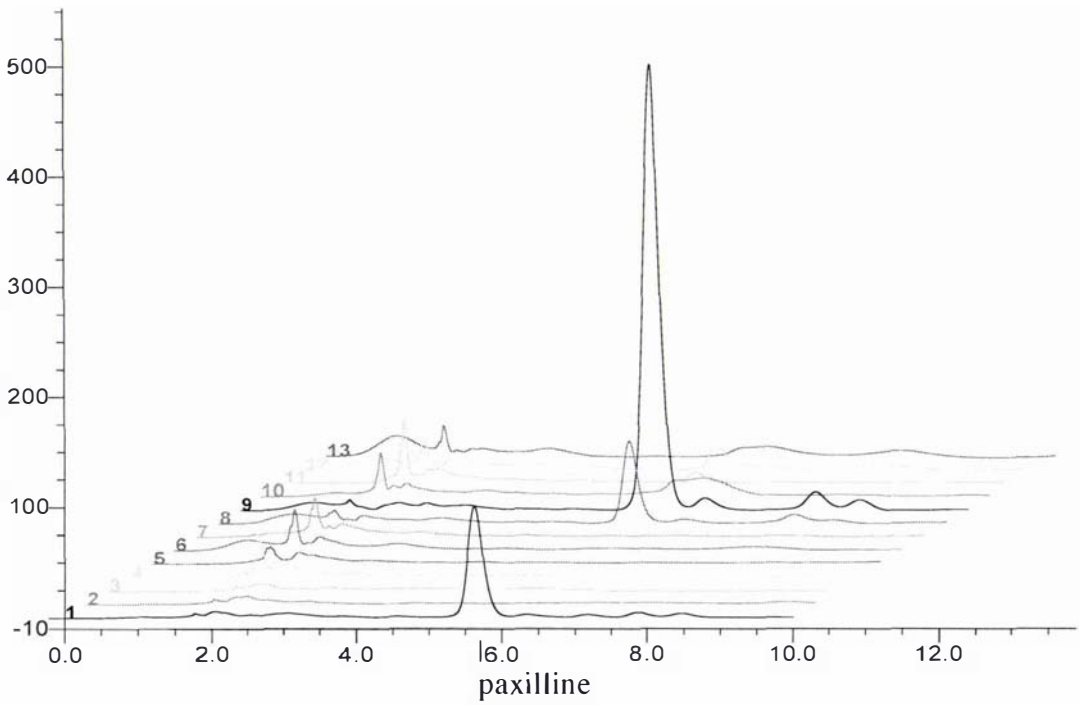
APPENDIX 5.4
HPLC ANALYSIS

5.5.1 HPLC analysis to confirm the TLC data in Fig 3.30
5.5.2 HPLC analysis to confirm the TLC data in Fig 3.50

Appendix 5.4.1 HPLC analysis to support Figure 3.30

RP-HPLC Analysis of *paxM* Complementations

Run by Sanjay Saikia
Solvent: 85:15 Methanol-Water (10 min run)
Injection volume: 10 µl
UV detection: 230 nm



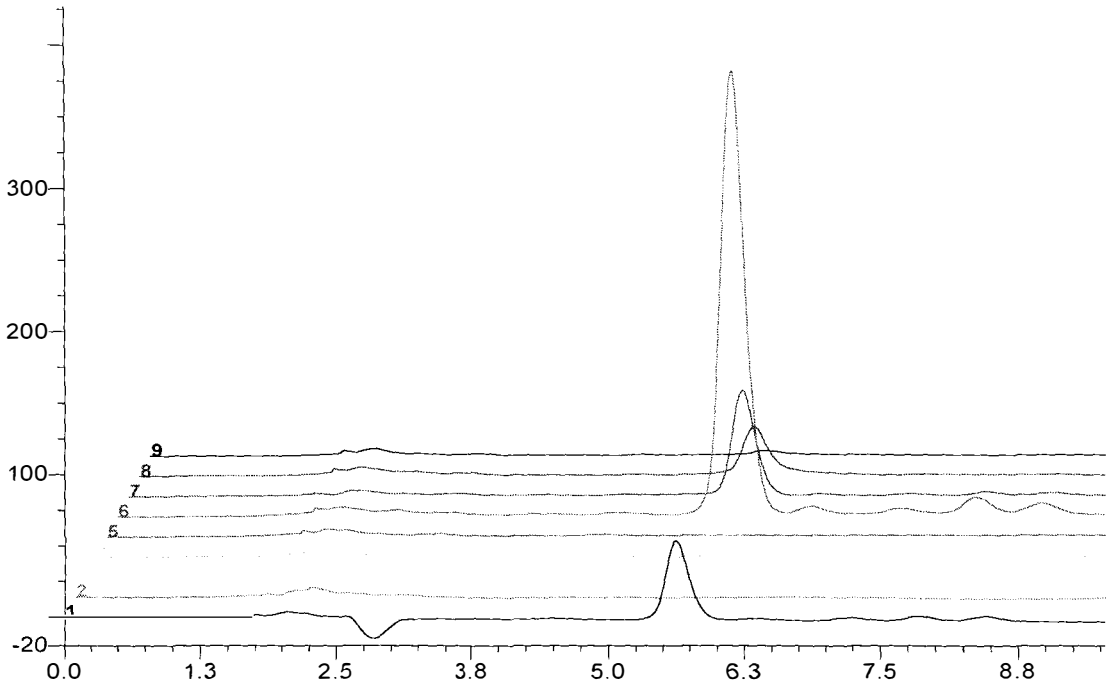
Trace 1: Wild type
Trace 2: LMM100
Trace 3: LMM200-1 (pII99)
Trace 4: LMM300-1 (pCY40)
Trace 5: LMM300-2 (pCY40)
Trace 6: LMM300-3 (pCY40)
Trace 7: LMM300-4 (pCY40)
Trace 8: LMM400-3 (pCY41)
Trace 9: LMM400-4 (pCY41)
Trace 10: LMM500-7 (pII99+pCY54)
Trace 11: LMM500-8 (pII99+pCY54)
Trace 12: LMM600-2 (pII99+pCY55)
Trace 13: LMM600-6 (pII99+pCY55)

Traces 1, 8, 9, 11, 12 and 13 are positive for paxilline (~5.6 min). In Traces 10 and 13 a clear peak for paxilline was not observed.

Appendix 5.4.2 HPLC analysis to support Figure 3.50

RP-HPLC Analysis of *paxC* Complementations

Run by Sanjay Saikia
Solvent: 85:15 Methanol-Water (10 min run)
Injection volume: 10 µl
UV detection: 230 nm



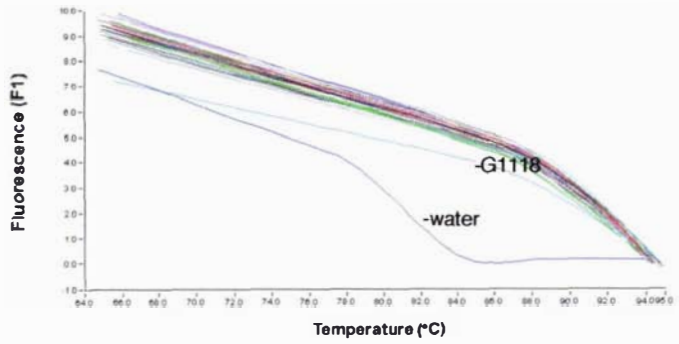
Trace 1: Wild type
Trace 2: ABC83
Trace 3: ABC283-1 (pII99)
Trace 4: ABC383-2 (pII99+pCY66)
Trace 5: ABC383-5 (pII99+pCY66)
Trace 6: ABC483-1 (pJA8)
Trace 7: ABC483-2 (pJA8)
Trace 8: ABC583-2 (pII99+pCY34)
Trace 9: ABC583-7 (pII99+pCY34)

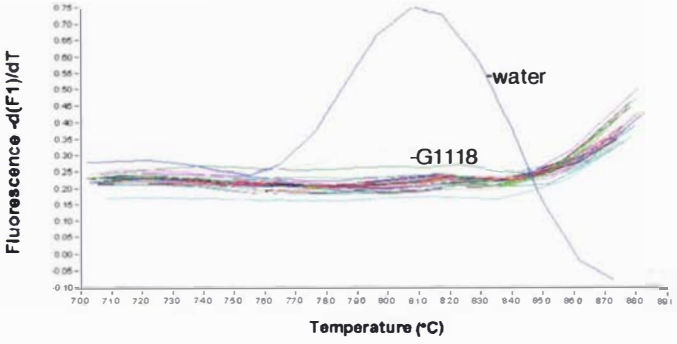
Traces 1 and 6 to 8 are positive for paxilline (~5.6 min).

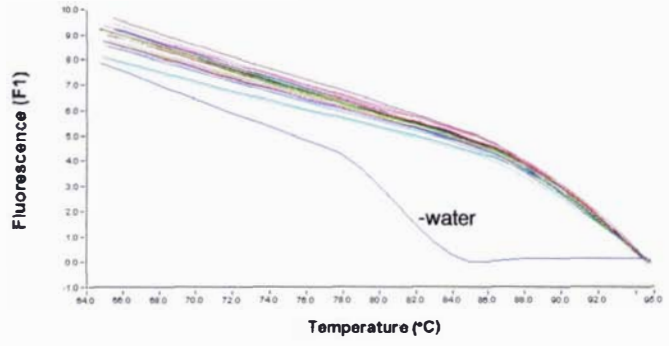
APPENDIX 5.5
MELTING CURVE DATA

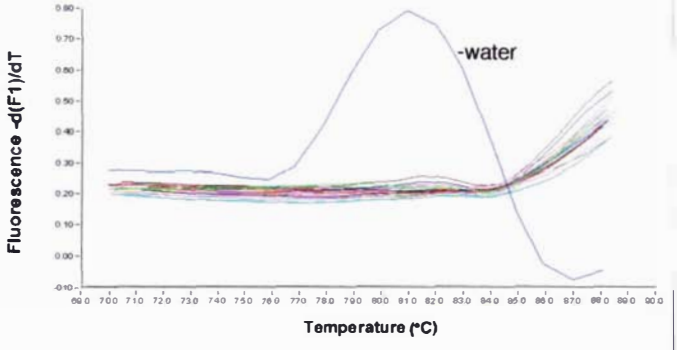
5.5.1 Melting curve analysis of the plant gene amplicon shown in Fig. 3.25
5.5.2 Melting curve analysis of the endophyte gene amplicon shown in Fig. 3.25

Appendix 5.5.1 Melting curves for the samples presented in Fig. 3.25A



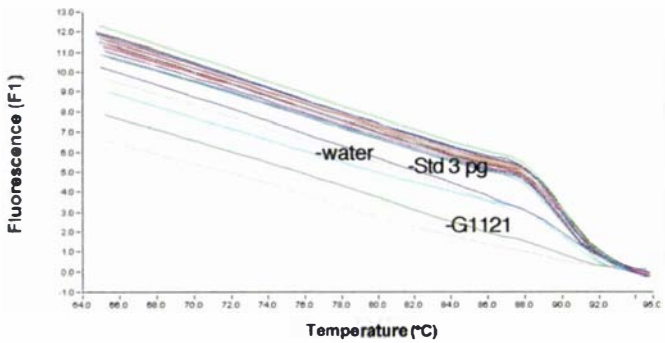


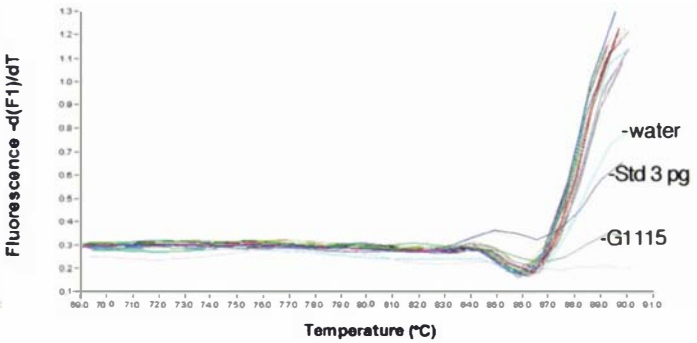


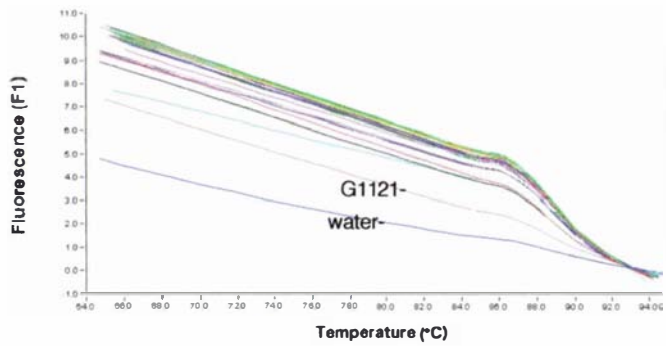


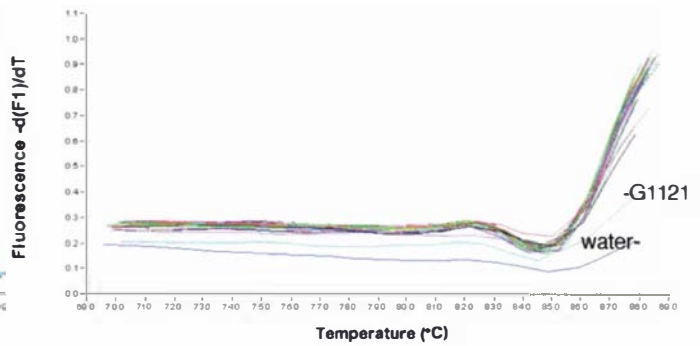
These samples have been amplified with primers prg12 and prg13 designed to the perennial ryegrass β-tubulin gene. The amplification of a product in the water sample is due to primer dimers. The G1118 sample was of low concentration and DNA quality.

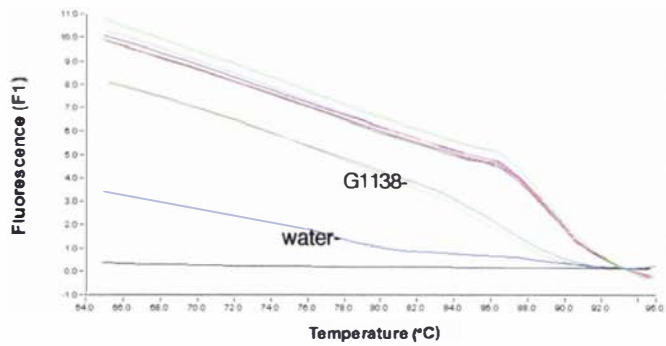
Appendix 5.5.2 Melting curves for the samples presented in Fig. 3.25B

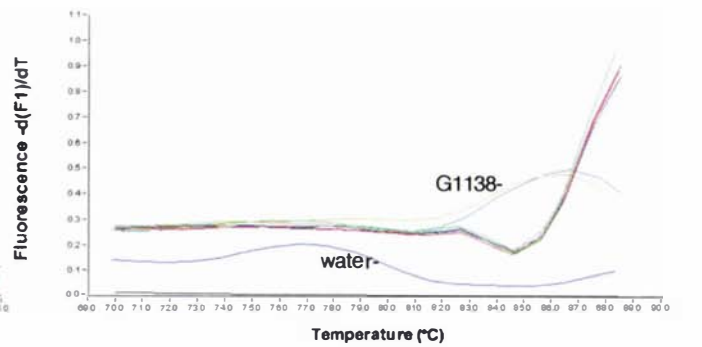












These samples have been amplified with primers endo1 and endo2 designed to the endophyte *ggs1* gene. The amplification of a product in the water sample is due to primer dimers. The samples G1115, G1121 and G1138 are endophyte free and amplification products in these samples are due to primer dimers.

References

Abe Y, Ono C, Hosobuchi M, Yoshikawa H (2002a) Functional analysis of *mlcR*, a regulatory gene for ML-236B (compactin) biosynthesis in *Penicillium citrinum*. Mol Genet Genomics 268: 352-361

Abe Y, Suzuki T, Mizuno T, Ono C, Iwamoto K, Hosobuchi M, Yoshikawa H (2002b) Effect of increased dosage of the ML-236B (compactin) biosynthetic gene cluster on ML-236B production in *Penicillium citrinum*. Mol Genet Genomics 268: 130-137

Abe Y, Suzuki T, Ono C, Iwamoto K, Hosobuchi M, Yoshikawa H (2002c) Molecular cloning and characterization of an ML-236B (compactin) biosynthetic gene cluster in *Penicillium citrinum*. Mol Genet Genomics 267: 636-646

Ahn J, Cheng YA, Walton JD (2002) An extended physical map of the Tox2 locus of *Cochliobolus carbonum* required for biosynthesis of HC-toxin. Fungal Genet Biol 35: 31-38

Ahn J, Walton JD (1998) Regulation of cyclic peptide biosynthesis and pathogenicity in *Cochliobolus carbonum* by TOXEp, a novel protein with a bZIP basic DNA-binding motif and four ankyrin repeats. Mol Gen Genet 260: 462-469

Ahn J, Walton JD (1997) A fattyacid synthase gene in *Cochliobolus carbonum* required for productionof HC-toxin, cyclo(D-prolyl-L-alanyl-D-alanyl-L-2-amino-9, 10-epoxi-8-oxodecanoyl). Mol Plant Microbe Interact 10: 207-214

Ahn J, Walton JD (1996) Chromosomal organisation of TOX2, a complex locus controlling host-selective toxin biosynthesis in *Cochliobolus carbonum*. Plant Cell 8: 887-897

Akamatsu H, Taga M, Kodama M, Johnson R, Otani H, Kohmoto K (1999) Molecular karyotypes for *Alternaria* plant pathogens known to produce host-specific toxins. Curr Genet 35: 647-656

Alexander NJ, Hohn TM, McCormick SP (1998) The TRI11 gene of *Fusarium sporotrichioides* encodes a cytochrome P-450 monooxygenase required for C-15 hydroxylation in trichothecene biosynthesis. Appl Environ Microbiol 64: 221-225

Alexander NJ, McCormick SP, Hohn TM (1999) TRI12, a trichothecene efflux pump from *Fusarium sporotrichioides*: gene isolation and expression in yeast. Mol Gen Genet 261: 977-984

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403-410

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-3402

Amnuaykanjanasin A, Epstein L (2003) A class V chitin synthase gene, *chsA* is essential for conidial and hyphal wall strength in the fungus *Colletotrichum graminicola* (*Glomerella graminicola*). Fungal Gen Biol 38: 272-285

Arntz C, Tudzynski P (1997) Identification of genes induced in alkaloid producing cultures of *Claviceps* sp. Curr Genet 31: 357-360

Barbato C, Calissano M, Pickford A, Romano N, Sandmann G, Macino G (1996) Mild RIP - an alternative method for in vivo mutagenesisof the albino-3 gene in *Neurospora crassa*. Mol Gen Genet 252: 353-361

Barredo JL, Diez B, Alvarez E, Martin JF (1989) Large amplification of a 35-kb DNA fragment carrying two penicillin biosynthetic genes in high penicillin producing strains of *Penicillium chrysogenum*. Curr Genet 16: 453-459.

Basse CW, Kolb S, Kahmann R (2002) A maize-specifically expressed gene cluster in *Ustilago maydis*. Mol Microbiol 43: 75-93

Basse CW, Stumpferl S, Kahmann R (2000) Characterization of a *Ustilago maydis* gene specifically induced during the biotrophic phase: evidence for negative as well as positive regulation. Mol Cell Biol 20: 329-339

Blankenship JD, Spiering MJ, Wilkinson HH, Fannin FF, Bush LP, Schardl CL (2001) Production of loline alkaloids by the grass endophyte, *Neotyphodium uncinatum*, in defined media. Phytochemistry 58: 395-401

Bohnert HU, Fudal I, Dioh W, Tharreau D, Notteghem J, Lebrun M (2004) A putative polyketide synthase/peptide synthetase from *Magnaporthe grisea* signals pathogen attack to resistant rice. Plant Cell 16: 2499-2513

Bojja RS, Cerny RL, Proctor RH, Du L (2004) Determining the biosynthetic sequence in the early steps of the fumonisin pathway by use of three gene-disruption mutants of *Fusarium verticillioides.* J Agric Food Chem 52: 2855-2860

Bok JW, Keller NP (2004) LaeA, a regulator of secondary metabolism in *Aspergillus* spp. Eukaryot Cell 3: 527-535

Bradshaw RE, Bhatnagar D, Ganley RJ, Gillman CJ, Monahan BJ, Seconi JM (2002) *Dothistroma pini*, a forest pathogen, contains homologs of aflatoxin biosynthetic pathway genes. Appl Environ Microbiol 68: 2885-2892

Braun CJ, Siedow JN, Williams ME, Levings III CS (1989) Mutations in the maize mitochondrial T-urf13 gene eliminate sensitivity to a fungal pathotoxin. Proc Nat Acad Sci USA 86: 4435-4439

Brown DW, Adams TH, Keller NP (1996) *Aspergillus* has distinct fatty acid synthases for primary and secondary metabolism. Proc Nat Acad Sci USA 93: 14873-14877

Brown DW, Dyer RB, McCormick SP, Kendra DF, Plattner RD (2004) Functional demarcation of the *Fusarium* core trichithecene gene cluster. Fungal Gen Biol 41: 454-462

Brown DW, McCormick SP, Alexander NJ, Proctor RH, Desjardins AE (2002) Inactivation of a cytochrome P-450 is a determinant of trichothecene diversity in *Fusarium* species. Fungal Genet Biol 36: 224-233

Brown DW, McCormick SP, Alexander NJ, Proctor RH, Desjardins AE (2001) A genetic and biochemical approach to study trichothecene diversity in *Fusarium sporotrichioides* and *Fusarium graminearum.* Fungal Genet Biol 32: 121-133

Brown DW, Proctor RH, Dyer RB, Plattner RD (2003) Characterization of a *Fusarium* 2-gene cluster involved in trichothecene C-8 modification. J Agric Food Chem 51: 7936-7944

Brown DW, Yu J-, Kelkar HS, Fernandes M, Nesbitt TC, Keller NP, Adams TH, Leonard TJ (1996) Twenty-five coregulated transcripts define a sterigmatocystin gene cluster in *Aspergillus nidulans*. Proc Natl Acad Sci USA 93: 1418-1422

Bruchez JJP, Eberle J, Russo VEA (1993) Regulatory sequences involved in the translation of *Neurospora crassa* mRNA: Kozak sequences and stop codons. Fungal Genet Newsl 40: 85-88

Bultman TL, White JF,Jr, Bowdish TI, Welch AM, Johnston J (1995) Mutualistic transfer of *Epichloë* spermatia by *Phorbia* flies. Mycologia 87: 182-189

Bush LP, Fannin FF, Siegel MR, Dahlman DL, Burton HR (1993) Chemistry, occurrence and biological effects of saturated pyrrolizidine alkaloids associated with endophyte-grass interactions. Agric, Ecosyst Environ 44: 81-102

Bush LP, Wilkinson HH, Schardl CL (1997) Bioprotective alkaloids of grass-fungal endophyte symbioses. Plant Physiol 114: 1-7

Byrd AD, Schardl CL, Songlin PJ, Mogen KL, Siegel MR (1990) The b-tubulin gene of *Epichloë typhina* from perennial ryegrass (*Lolium perenne*). Curr Genet 18: 347-354

Cambareri EB, Jensen BC, Schabtach E, Selker EU (1989) Repeat-induced G-C to A-T mutations in *Neurospora*. Science 244: 1571-1575

Carroll AM, Sweigard JA, Valent B (1994) Improved vectors for selecting resistance to hygromycin. Fungal Genet Newsl 22

Cary JW, Wright M, Bhatnagar D, Lee R, Chu FS (1996) Molecular characterisation of an *Aspergillus parasiticus* dehydrogenase gene, *norA*, located on the aflatoxin biosynthesis gene cluster. Appl Environ Microbiol 62: 360-366

Chang PK (2003) The *Aspergillus parasiticus* protein AFLJ interacts with the aflatoxin pathway-specific regulator AFLR. Mol Genet Genomics 268: 711-719

Chang P, Cary J, W., Bhatnagar D, Cleveland TE, Bennett JW, Linz JE, Woloshuk CP, Payne GA (1993) Cloning of the *Aspergillus parasiticus apa-2* gene associated with the regulation of aflatoxin biosynthesis. Appl Environ Microbiol 59: 3273-3279

Chang P, Yu J (2002) Characterization of a partial duplication of the aflatoxin gene cluster in *Aspergillus parasiticus* ATCC 56775. Appl Microbiol Biotechnol 58: 632-636

Chang PK, Yu J, Ehrlich KC, Boue SM, Montalbano BG, Bhatnagar D, Cleveland TE (2000) *adhA* in *Aspergillus parasiticus* is involved in conversion of 5'-hydroxyaverantin to averufin. Appl Environ Microbiol 66: 4715-4719

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

Chen AP, Kroon PA, Poulter CD (1994) Isoprenyl diphosphate synthases: protein sequence comparisons, a phylogenetic tree, and predictions of secondary structure. Protein Science 3: 600-607

Cheng Y, Ahn J-, Walton JD (1999) A putative branched-chain-amino-acid transaminase gene required for HC-toxin biosynthesis and pathogenicity in *Cochliobolus carbonum*. Microbiology 145: 3539-3546

Cheng Y, Walton JD (2000) A eukaryotic alanine racemase gene involved in cyclic peptide biosynthesis. J Biol Chem 275: 4906-4911

Chiou C-, Miller M, Wilson DL, Trail F, Linz JE (2002) Chromosomal location plays a role in regulation of aflatoxin gene expression in *Aspergillus parasiticus*. Appl Environ Microbiol 68: 306-315

Christensen MJ, Bennett RJ, Schmid J (2002) Growth of *Epichloë/Neotyphodium* and p-endophytes in leaves of *Lolium* and *Festuca* grasses. Mycol Res 106: 93-106

Christensen MJ, Leuchtmann A, Rowan DD, Tapper BA (1993) Taxonomy of Acremonium endophytes of tall fescue (*Festuca arundinacea*), meadow fescue (*F. pratensis*) and perennial rye-grass (*Lolium perenne*). Mycol Res 97: 1083-1092

Chung K, Leuchtmann A, Schardl CL (1996) Inheritance of mitochondrial DNA and plasmids in the Ascomycetous fungus, *Epichloë typhina*. Genetics 142: 259-265

Chung K, Schardl CL (1997) Vegetative compatibility between and within *Epichloë* species. Mycologia 89: 558-565

Clay K (1990) Fungal endophytes of grasses. Annu Rev Ecol Syst 21: 275-297

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

Cole RJ, Cox RH (1981) Handbook of toxic fungal metabolites. Academic Press, London

Cole RJ, Dorner JW, Lansden JA, Cox RH, Pape C, Cunfer BM, Nicholson SS, Bendell DM (1977) Paspalum staggers: isolation and identification of tremorgenic metabolites from sclerotia of *Claviceps paspali*. J Agric Food Chem 25: 1197-1201

Cole RJ, Kirksey JW, Wells JM (1974) A new tremorgenic metabolite from *Penicillium paxilli*. Can J Microbiol 20: 1159-1162

Collett MA, Bradshaw RE, Scott DB (1995) A mutualistic fungal symbiont of perennial ryegrass contains two different pyr4 genes, both expressing orotidine-5'-monophosphate decarboxylase. Gene 158: 31-39

Correia T, Grammel N, Ortel I, Keller U, Tudzynski P (2003) Molecular cloning and analysis of the ergopeptine assembly system in the ergot fungus *Claviceps purpurea*. Chem Biol 10: 1281-1292

Cox GB, Stout RW (1987) Study of the retention mechanism for basic compounds on silica under 'pseudo-reveresed phase' conditions. J Chromatogr 384: 315-336

Daboussi M (1997) Fungal transposable elements and genome evolution. Genetica 100: 253-260

Desjardins AE, Plattner RD, Proctor RH (1996) Linkage among genes responsible for fumonisin biosynthesis in *Gibberella fujikuroi* mating population A. Appl Environ Microbiol 62: 2571-2576

Diez B, Gutierrez S, Barredo JL, van Solingen P, van der Voort LH, Martin JF (1990) The cluster of penicillin biosynthetic genes. Identification and characterization of the *pcbAB* gene encoding the alpha-aminoadipyl-cysteinyl-valine synthetase and linkage to the *pcbC* and *penDE* genes. J Biol Chem 265: 16358-16365

Dobson JM (1997) Molecular characterisation of the HMG CoA reductase gene from *Neotyphodium lolii*. PhD thesis. Massey University.

Dorner JW, Cole RJ, Cox RH, Cunfer BM (1984) Paspalitrem C, a new metabolite from sclerotia of *Claviceps paspali*. J Agric Food Chem 32: 1069-1071

Dower WJ, Miller JF, Ragsdale CW (1988) High efficiency transformation of *E. coli* by high voltage electroporation. Nucleic Acids Res 16: 6127-6145

Ehrlich KC, Chang PK, Yu J, Cotty PJ (2004) Aflatoxin biosynthesis cluster gene *cypA* is required for G aflatoxin formation. Appl Environ Microbiol 70: 6518-6524

Farfsing JW, Auffarth K, Basse CW (2005) Identification of cis-Active elements in *Ustilago maydis mig2* promoter conferring high-level activity during pathogenic growth in maize. Mol Plant Microbe Interact 18: 75-87

Felitti S, Shields K, Ramsperger M, Tian P, Webster T, Ong B, Sawbridge T, Spangenberg G (2003) Gene discovery and microarray-based transcriptome analysis in grass endophytes. Kluwer Academic, Dordrecht, The Netherlands

Feng GH, Leonard TJ (1995) Characterization of the polyketide synthase gene (*pksL1*) required for aflatoxin biosynthesis in *Aspergillus parasiticus*. J Bacteriol 177: 6246-6254

Fernandes M, Keller NP, Adams TH (1998) Sequence-specific binding by *Aspergillus nidulans* AflR, a C6 zinc cluster protein regulating mycotoxin biosynthesis. Mol Microbiol 28: 1355-1365

Fierro F, Barredo JL, Díez B, Gutierrez S, Fernández FJ, Martín JF (1995) The penicillin gene cluster is amplified in tandem repeats linked by sonserved hexanucleotide sequences. Proc Natl Acad Sci USA 92: 6200-6204

Fierro F, Gurtiérrez S, Díez B, Martín JF (1993) Resolution of four large chromosomes in penicillin-producing filamentous fungi: the penicillin gene cluster is located on chromosome II (9.6 Mb) in *Penicillium notatum* and chromosome I (10.4 Mb) in *Penicillium chrysogenum*. Mol Gen Genet 241: 573-578

Flaherty JE, Pirttila AM, Bluhm BH, Woloshuk CP (2003) PAC1, a pH-regulatory gene from *Fusarium verticillioides*. Appl Environ Microbiol 69: 5222-5227

Flaherty JE, Woloshuk CP (2004) Regulation of fumonisin biosynthesis in *Fusarium verticillioides* by a zinc binuclear cluster-type gene, *ZFR1*. Appl Environ Microbiol 70: 2653-2659

Fletcher LR (1999) "Non-toxic" endophytes in ryegrass and their effect on livestock health and production. In DR Woodfield, C Matthew, eds, Ryegrass endophyte: an essential New Zealand symbiosis. New Zealand Grassland Association, Napier, New Zealand, pp 133-139

Fueki S, Tokiwano T, Toshima H, Oikawa H (2004) Biosynthesis of indole diterpenes, emindole, and paxilline: involvement of a common intermediate. Org Lett 6: 2697-2700

Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma LJ, Smirnov S, Purcell S, Rehman B, Elkins T, Engels R, Wang S, Nielsen CB, Butler J, Endrizzi M, Qui D, Ianakiev P, Bell-Pedersen D, Nelson MA, Werner-Washburne M, Selitrennikoff CP, Kinsey JA, Braun EL, Zelter A, Schulte U, Kothe GO, Jedd G, Mewes W, Staben C, Marcotte E, Greenberg D, Roy A, Foley K, Naylor J, Stange-Thomann N, Barrett R, Gnerre S, Kamal M, Kamvysselis M, Mauceli E, Bielke C, Rudd S, Frishman D, Krystofova S, Rasmussen C, Metzenberg RL, Perkins DD, Kroken S, Cogoni C, Macino G, Catcheside D, Li W, Pratt RJ, Osmani SA, DeSouza CP, Glass L, Orbach MJ, Berglund JA, Voelker R, Yarden O, Plamann M, Seiler S, Dunlap J, Radford A, Aramayo R, Natvig DO, Alex LA, Mannhaupt G, Ebbole DJ, Freitag M, Paulsen I, Sachs MS, Lander ES, Nusbaum C, B. B (2003) The genome sequence of the filamentous fungus *Neurospora crassa*. Nature 422: 859-868

Gallagher RT, Hawkes AD, Steyn PS, Vleggaar R (1984) Tremorgenic neurotoxins from perennial ryegrass causing ryegrass staggers disorder of livestock: structure elucidation of lolitrem B. J Chem Soc Chemical Communications 614-616

Gallagher RT, Wilson BJ (1978) Aflatrem, a tremorgenic toxin from *Aspergillus flavus*. Mycopathologia 66: 183-185

Ganley ARD, Scott B (2002) Concerted evolution in the ribosomal RNA genes of an *Epichloë* endophyte hybrid: comparison between tandemly-arranged rDNA and dispersed 5S rrn genes. Fungal Genet Biol 35: 39-51

Ganley ARD, Scott B (1998) Extraordinary ribosomal spacer length heterogeneity in a *Neotyphodium* endophyte hybrid: implications for concerted evolution. Genetics 150: 1625-1637

Gardiner DM, Cozijnsen AJ, Wilson LM, Pedras MS, Howlett BJ (2004) The sirodesmin biosynthetic gene cluster of the plant pathogenic fungus *Leptosphaeria maculans*. Mol Microbiol 53: 1307-1318

Gardiner DM, Jarvis RS, Howlett BJ (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 Genet Biol 42: 257-263

Gatenby WA, Munday-Finch SC, Wilkins AL, Miles CO (1999) Terpendole M, a novel indole-diterpenoid isolated from *Lolium perenne* infected with the endophytic fungus *Neotyphodium lolii*. J Agric Food Chem 47: 1092-1097

Gebler JC, Poulter CD (1992) Purification and characterisation of dimethylallyl tryptophan synthase from *Claviceps purpurea*. Arch Biochem Biophys 296: 308-313

Gebler JC, Woodside AB, Poulter CD (1992) Dimethylallyltryptophan synthase. An enzyme-catalyzed electrophilic aromatic substitution. J Am Chem Soc 114: 7354-7360

Glenn AE, Bacon CW, Price R, Hanlin RT (1996) Molecular phylogeny of *Acremonium* and its taxonomic implications. Mycologia 88: 369-383

Gloer JB (1995) Antiinsectan natural products from fungal sclerotia. Acc Chem Res 28: 343-350

González MC, Lull C, Moya P, Ayala I, Yúfera PJ (2003) Insecticidal activity of penitrems, including penitrem G, a new member of the family isolated from *Penicillium crustosum*. J Agric Food Chem 51: 2156-2160

Graïa F, Lespinet O, Rimbault B, Dequard-Chablat M, Coppin E, Picard M (2001) Genome quality control: RIP (repeat-induced point mutation) comes to *Podospora*. Mol Microbiol 40: 586-595.

Guzmán-de-Peña D, Ruiz-Herrera J (1997) Relationship between aflatoxin biosynthesis and sporulation in *Aspergillus parasiticus*. Fungal Genet Biol 21: 198-205

Gwinn KD, Collins-Shepard MH, Reddick BB (1991) Tissue print-immunoblot, an accurate method for the detection of *Acremonium coenophialum* in tall fescue. Phytopathology 81: 747-748

Hamann A, Feller F, Osiewacz HD (2000) The degenerate DNA transposon Pat and repeat-induced point mutation (RIP) in *Podospora anserina*. Mol Gen Genet 263: 1061-1069

Hamer L, Adachi K, Montenegro-Chamorro MV, Tanzer MM, Mahanty SK, Lo C, Tarpey RW, Skalchunes AR, Heiniger RW, Frank SA, Darveaux BA, Lampe DJ, Slater TM, Ramamurthy L, DeZwaan TM, Nelson GH, Shuster JR, Woessner J, Hamer JE (2001) Gene discovery and gene function assignment in filamentous fungi. Proc Natl Acad Sci U S A 98: 5110-5115

Han S, Adams TH (2001) Complex control of the developmental regulatory locus *brlA* in *Aspergillus nidulans*. Mol Genet Genomics 266: 260-270

Han Y, Liu X, Benny U, Kistler HC, VanEtten HD (2001) Genes determining pathogenicity to pea are clustered on a supernumerary chromosome in the fungal plant pathogen *Nectria haematococca*. Plant J 25: 305-314

Harris LJ, Desjardins AE, Plattner RD, Nicholson P, Butler G, Young JC, Weston G, Proctor RH, Hohn TM (1999) Possible role of trichothecene mycotoxins in virulence of *Fusarium graminearum* on maize. Plant Dis 83: 954-960

Hatta R, Ito K, Tanaka T, Tanaka A, Yamamoto M, Akimitsu K, Tsuge T (2002) A conditionally dispensable chromosome controls host-specific pathogenicity in the fungal plant pathogen *Alternaria alternata*. Genetics 161: 59-70

Hedden P, Phillips AL, Rojas MC, Carrera E, Tudzynski B (2002) Gibberellin biosynthesis in plants and fungi: A case of convergent evolution? J Plant Growth Regul 20: 319-331

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

Hicks JK, Yu J-, Keller NP, Adams TH (1997) *Aspergillus* sporulation and mycotoxin production both require inactivation of the FadA Ga protein-dependent signaling pathway. EMBO J 16: 4916-4923

Hohn TM, Beremand PD (1989) Isolation and nucleotide sequence of a sesquiterpene cyclase gene from the trichothecene-producing fungus *Fusarium sporotrichiodes*. Gene 79: 131-138

Hohn TM, Desjardins AE (1992) Isolation and gene disruption of the *Tox5* gene encoding trichodiene synthase in *Gibberella pulicaris*. Mol Plant Microbe Interact 5: 214-222

Hohn TM, Desjardins AE, McCormick SP (1995) The *Tri4* gene of *Fusarium sporotrichioides* encodes a cytochrome P450 monooxygenase involved in trichothecene biosynthesis. Mol Gen Genet 248: 95-102

Hohn TM, Desjardins AE, McCormick SP (1993) Analysis of *Tox5* gene expressionin *Gibberella pulicaris* strains with different trichothecene production phenotypes. Appl Environ Microbiol 59: 2359-2363

Hohn TM, Krishna R, Proctor RH (1999) Characterization of a transcriptional activator controlling trichothecene toxin biosynthesis. Fungal Genet Biol 26: 224-235

Hohn TM, McCormick SP, Desjardins AE (1993) Evidence for a gene cluster involving trichothecene-pathway biosynthetic genes in *Fusarium sporotrichioides*. Curr Genet 24: 291-295

Horiuchi H, Fujiwara M, Yamashita S, Ohta A, Takagi M (1999) Proliferation of intrahyphal hyphae caused by disruption of *csmA*, which encodes a class V chitin synthase with a myosin motor-like domain in *Aspergillus nidulans*. J Bacteriol 181: 3721-3729

Hua-Van A, Hericourt F, Capy P, Daboussi MJ, Langin T (1998) Three highly divergent subfamilies of the impala transposable element coexist in the genome of the fungus *Fusarium oxysporum*. Mol Gen Genet 259: 354-362

Huson DH (1998) SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics 14:68-73

Idnurm A, Howlett BJ (2003) Analysis of loss of pathogenicity mutants reveals that repeat-induced point mutations can occur in the Dothideomycete *Leptosphaeria maculans*. Fungal Genet Biol 39: 31-37

Ito H, Tanaka T, Hatta R, Yamamoto M, Akimitsu K, Tsuge T (2004) Dissection of the host range of the fungal plant pathogen *Alternaria alternata* by modification of secondary metabolism. Mol Microbiol 52: 399-411

Itoh Y, Johnson R, Scott B (1994) Integrative transformation of the mycotoxin-producing fungus, *Penicillium paxilli*. Curr Genet 25: 508-513

Jiang Y, Proteau P, Poulter D, Ferro-Novick S (1995) BTS1 encodes a geranylgeranyl diphosphate synthase in Saccharomyces cerevisiae. J Biol Chem 270: 21793-21799

Johnson LJ, Johnson RD, Akamatsu H, Salamiah A, Otani H, Kohmoto K, Kodama M (2001) Spontaneous loss of a conditionally dispensable chromosome from the *Alternaria alternata* apple pathotype leads to loss of toxin production and pathogenicity. Curr Genet 40: 65-72

Johnson RD, Johnson L, Itoh Y, Kodama M, Otani H, Kohmoto K (2000) Cloning and characterization of a cyclic peptide synthetase gene from *Alternaria alternata* apple pathotype whose product is involved in AM-toxin synthesis and pathogenicity. Mol Plant Microbe Interact 13: 742-753

Kelkar HS, Skloss TW, Haw JF, Keller NP, Adams TH (1997) *Aspergillus nidulans stcL* encodes a putative cytochrome P-450 monooxygenase required for bisfuran desaturation furing aflatoxin/sterigmatocystin biosynthesis. J Biol Chem 272: 1589-1594

Keller NP, Hohn TM (1997) Metabolic pathway gene clusters in filamentous fungi. Fungal Genet Biol 21: 17-29

Keller NP, Kantz NJ, Adams TH (1994) *Aspergillus nidulans verA* is required for production of the mycotoxin sterigmatocystin. Appl Environ Microbiol 60: 1444-1450

Keller NP, Watanabe CMH, Kelkar HS, Adams TH, Townsend CA (2000) Requirement of monooxygenase-mediated steps for sterigmatocystin biosynthesis by *Aspergillus nidulans*. Appl Environ Microbiol 66: 359-362.

Kellis M, Birren BW, Lander ES (2004) Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. Nature 428: 617-624

Kellis M, Patterson N, Endrizzi M, Birren B, Lander ES (2003) Sequencing and comparison of yeast species to identify genes and regulatory elements. Nature 423: 241-254

Kempken F, Kück U (1998) Transposons in filamentous fungi-facts and perspectives. Bioessays 20: 652-659

Kennedy J, Auclair K, Kendrew SG, Park C, Vederas JC, Hutchinson CR (1999) Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. Science 284: 1368-1372

Kim JM, Vanguri S, Boeke JD, Gabriel A, Voytas DF (1998) Transposable elements and genome organization: A comprehensive survey of retrotransposons revealed by the complete *Saccharomyces cerevisiae* genome sequence. Genome Res 8: 464-478

Kimura M, Kaneko I, Komiyama M, Takatsuki A, Koshino H, Yoneyama K, Yamaguchi I (1998a) Trichothecene 3-O-acetyltransferase protects both the producing organism and transformed yeast from related mycotoxins. Cloning and characterization of *Tri101*. J Biol Chem 273: 1654-1661

Kimura M, Matsumoto G, Shingu Y, Yoneyama K, Yamaguchi I (1998b) The mystery of the trichothecene 3-O-acetyltransferase gene. FEBS Lett 435: 163-168

Kimura M, Tokai T, O'Donnell K, Ward TJ, Fujimura M, Hamamoto H, Shibata T, Yamaguchi I (2003) The trichothecene biosynthesis gene cluster of *Fusarium graminearum* F15 contains a limited number of essential pathway genes and expressed non-essential genes. FEBS Lett 539: 105-110

Knaus H-, McManus OB, Lee SH, Schmalhofer WA, Garcia-Calvo M, Helms LMH, Sanchez M, Giangiacomo K, Reuben JP, Smith AB, Kaczorowski GJ, Garcia ML (1994) Tremorgenic indole alkaloids potently inhibit smooth muscle high-conductance calcium-activated channels. Biochemistry 33: 5819-5828

Kodama M, Rose MS, Yang G, Yun SH, Yoder OC, Turgeon BG (1999) The translocation-associated Tox1 locus of *Cochliobolus heterostrophus* is two genetic elements on two different chromosomes. Genetics 151: 585-596

Kroken S, Glass NL, Taylor JW, Yoder OC, Turgeon BG (2003) Phylogenomic analysis of type I polyketide synthase genes in pathogenic and sprobic ascomycetes. Proc Natl Acad Sci USA 100: 15672-15675

Kuldau GA, Tsai H-, Schardl CL (1999) Genome sizes of *Epichloë* species and anamorphic hybrids. Mycologia 91: 776-782

Kusumoto K-, Nogata Y, Ohta H (2000) Directed deletions in the aflatoxin biosynthesis gene homolog cluster of *Aspergillus oryae*. Curr Genet 37: 104-111

Kusumoto K, Yabe K, Nogata Y, Ohta H (1998) Transcript of a homolog of *aflR*, a regulatory gene for aflatoxin synthesis in *Aspergillus parasiticus*, was not detected in *Aspergillus oryae* strains. FEBS Microbiol Letters 169: 303-307

Kutil BL, Liu G, Vrebalov J, Wilkinson HH (2004) Contig assembly and microsynteny analysis using a bacterial artificial chromosome library for *Epichloë festucae*, a mutualistic fungal endophyte of grasses. Fungal gen Biol 41: 23-32

Lane GA, Christensen MJ, Miles CO (2000) Coevolution of fungal endophytes with grasses: the significance of secondary metabolites. In CW Bacon, JFJ White, eds, Microbial Endophytes. Marcel Dekker, New York, pp 341-388

Latch GCM, Christensen MJ (1985) Artificial infection of grasses with endophytes. Ann Appl Biol 107: 17-24

Latch GCM, Christensen MJ, Tapper BA, Easton HS, Hume DE, Fletcher LR (2000) Ryegrass endophytes. Patent.

Leblanc P, Desset S, Dastugue B, Vury C (1997) Invertebrate retroviruses: ZAM a new candidate in D. melanogaster. EMBO J 16: 7521-7531

Lee T, Han Y-, Kim K-, Yun S-, Lee Y- (2002) Tri13 and Tri7 determine deoxynivalenol- and nivalenol-producing chemotypes of *Gibberella zeae*. Appl Environ Microbiol 68: 2148-2154

Lindstrom J, T., Sun S, Belanger F, C. (1993) A novel fungal protease expressed in endophytic infection of *Poa* species. Plant Physiol 102: 645-650

Lindstrom JT, Belanger FC (1994) Purification and characterisation of an endophytic fungal proteinase that is abundantly expressed in the infected host grass. Plant Physiol 106: 7-16

Lu S, Lyngholm L, Yang G, Bronson C, Yoder OC, Turgeon BG (1994) Tagged mutations at the Tox1 locus of *Cochliobolus heterostrophus* by restriction enzyme-mediated integration. Proc Natl Acad Sci USA 91: 12649-12653

MacCabe AP, Riach MBR, Unkles SE, Kinghorn JR (1990) The *Aspergillus nidulans npeA* locus consists of three contiguous genes required for penicillin biosynthesis. EMBO J 9: 279-287.

Madrid MP, Di Pietro A, Roncero MIG (2003) Class V chitin synthase determines pathogenesis in the vascular wilt fungus *Fusarium oxysporum* and mediates resistance to plant defence compounds. Mol Microbiol 47: 257-266

Majewska-Sawka A, Nakashima H (2004) Endophyte transmission via seeds of *Lolium perenne* L.: immunodection of fungal antigens. Fungal Gen Biol 41: 534-541

Malonek S, Rojas MC, Hedden P, Gaskin P, Hopkins P, Tudzynski B (2004) The NADPH: cytochrome P450 reductase gene from *Gibberella fujikuroi* is essential for gibberellin biosynthesis. J Biol Chem 279: 25075-25084

Maniatis T, Sambrook J, Fritsch EF (1982) Molecular Cloning - A Laboratory Manual. Cold Spring Harbor Press, Cold Spring Harbor, Cold Spring Harbor, New York

Mantle PG, Weedon CM (1994) Biosynthesis and transformation of tremorgenic indole-diterpenoids by *Penicillium paxilli* and *Acremonium lolii*. Phytochemistry 36: 1209-1217

Marahiel MA, Stachelhaus T, Mootz HD (1997) Modular peptide synthetases involved in nonribosomal peptide synthesis. Chem Rev 97: 2651-2673

McCormick SP, Alexander NJ (2002) *Fusarium* Tri8 encodes a trichothecene C-3 esterase. Appl Environ Microbiol 68: 2959-2964

McCormick SP, Alexander NJ, Trapp SE, Hohn TM (1999) Disruption of *TRI101*, the gene encoding trichothecene 3-O-acetyltransferase from *Fusarium sporotrichioides*. Appl Environ Microbiol 65: 5252-5256

McCormick SP, Hohn TM, Desjardins AE (1996) Isolation and characterization of *Tri3*, a gene encoding 15-O-acetyltransferase from *Fusarium sporotrichioides*. Appl Environ Microbiol 62: 353-359

McGill MK (2000) Cloning and characterisation of two subtilisin-like protease genes from *Neotyphodium lolii*. MSc thesis. Massey University.

McMillan LK, Carr RL, Young CA, Astin JW, Lowe RGT, Parker EJ, Jameson GB, Finch SC, Miles CO, McManus OB, Schmalhofer WA, Garcia ML, Kaczorowski GJ, Goetz MA, Tkacz JS, Scott B (2003) Molecular analysis of two cytochrome P450 monooxygenase genes required for paxilline biosynthesis in *Penicillium paxilli* and effects of paxilline intermediates on mammalian maxi-K ion channels. Mol Genet Genomics 270: 9-23

Meek IB, Peplow AW, Ake C,Jr, Phillips TD, Beremand MN (2003) *Tri1* encodes the cytochrome P450 monooxygenase for C-8 hydroxylation during trichothecene biosynthesis in *Fusarium sporotrichioides* and resides upstream of another new *Tri* gene. Appl Environ Microbiol 69: 1607-1613

Meeley RB, Johal GS, Briggs SP, Walton JD (1992) A biochemical phenotype for a disease resistace gene of maize. Plant Cell 4: 71-77

Mende K, Homann V, Tudzynski B (1997) The geranylgeranyl diphosphate synthase gene of *Gibberella fujikuroi*: isolation and expression. Mol Gen Genet 255: 96-105

Meyers DM, OBrian GR, Du WL, Bhatnagar D, Payne GA (1998) Characterisation of aflJ, a gene required for conversion of pathway intermediated to aflatoxin. Appl Environ Microbiol 64: 3713-3717

Mihlan M, Homann V, Tudzynski B (2003) AREA directly mediates nitrogen regulation of gibberellin biosynthesis in *Gibberella fujikuroi*, but its activity is not affected by NMR. Mol Microbiol 47: 975-991

Miles CO, di Menna ME, Jacobs SWL, Garthwaite I, Lane GA, Prestidge RA, Marshall SL, Wilkinson HH, Schardl CL, Ball OJ-, Latch GCM (1998) Endophytic fungi in indigenous Australasian grasses associated with toxicity to livestock. Appl Environ Microbiol 64: 601-606

Miles CO, Munday SC, Wilkins AL, Ede RM, Towers NR (1994) Large-scale isolation of lolitrem B and structure determination of lolitrem E. J Agric Food Chem 42: 1488-1492

Miller JH (1972) Experiments in Molecular Genetics. Cold Spring Harbor Laboratory Press, New York

Mogen KL, Siegel MR, Schardl CL (1991) Linear DNA plasmids of the perennial ryegrass choke pathogen, *Epichloë typhina* (Clavicipitaceae). Curr Genet 20: 519-526

Möller EM, Bahnweg G, Sandermann H, Geiger HH (1992) A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues. Nucleic Acids Res 20: 6115-6116.

Moon CD, Craven KD, Leuchtmann A, Clement SL, Schardl CL (2004) Prevalence of interspecific hybrids amongst asexual fungal endophytes of grasses. Mol Ecol 13: 1455-1467

Moon CD, Scott B, Schardl CL, Christensen MJ (2000) The evolutionary origins of *Epichloë* endophytes from annual ryegrasses. Mycologia 92: 1103-1118

Moon CD, Tapper BA, Scott B (1999) Identification of *Epichloë* endophytes *in planta* by a microsatellite-based PCR fingerprinting assay with automated analysis. Appl Environ Microbiol 65: 1268-1279

Moy M, Li HM, Sullivan R, White JF,Jr, Belanger FC (2002) Endophytic fungal beta-1,6-glucanase expression in the infected host grass. Plant Physiol 130: 1298-1308

Munday-Finch SC, Miles CO, Wilkins AL, Hawkes AD (1995) Isolation and structure elucidation of lolitrem A, a tremorgenic mycotoxin from perennial ryegrass infected with *Acremonium lolii*. J Agric Food Chem 43: 1283-1288

Munday-Finch SC, Wilkins AL, Miles CO (1998) Isolation of lolicine A, lolicine B, lolitriol, and lolitrem N from *Lolium perenne* infected with *Neotyphodium lolii* and evidence for the natural occurrence of 31-epilolitrem N and 31-epilolitrem F. J Agric Food Chem 46: 590-598

Munday-Finch SC, Wilkins AL, Miles CO (1996a) Isolation of paspaline B, an indole-diterpenoid from *Penicillium paxilli*. Phytochemisty 41: 327-332

Munday-Finch SC, Wilkins AL, Miles CO, Ede RM, Thomson RA (1996b) Structure elucidation of Lolitrem F, a naturally occurring stereoisomer of the tremorgenic mycotoxin Lolitrem B, isolated from *Lolium perenne* infected with *Acremonium lolii*. J Agric Food Chem 44: 2782-2788

Munday-Finch SC, Wilkins AL, Miles CO, Tomoda H, Omura S (1997) Isolation and structure elucidation of lolilline, a possible biosynthetic precursor of the lolitrem family of tremorgenic mycotoxins. J Agric Food Chem 45: 199-204

Murray FR, Latch GCM, Scott DB (1992) Surrogate transformation of perennial ryegrass, *Lolium perenne*, using genetically modified *Acremonium* endophyte. Mol Gen Genet 233: 1-9

Nakayashiki H, Nishimoto N, Ikeda K, Tosa Y, Mayama S (1999) Degenerate MAGGY elements in a subgroup of *Pyricularia grisea*: a possible example of successful capture of a genetic invader by a fungal genome. Mol Gen Genet 261: 958-966

Namiki F, Matsunaga M, Okuda M, Inoue I, Nishi K, Fujita Y, Tsuge T (2001) Mutation of an arginine biosynthesis gene causes reduced pathogenicity in *Fusarium oxysporum* f. sp. melonis. Mol Plant Microbe Interact 14: 580-584.

Nozawa K, Udagawa S, Nakajima S, Kawai K (1987) Structures of two stereoisomers of a new type of indoloditerpene related to the tremorgenic mycotoxin paxilline, from *Emericella desertorum* and *Emericella striata*. J Chem Soc Chem Commun 1157-1159

Okada K, Saito T, Nakagawa T, Kawamukai M, Kamiya Y (2000) Five geranylgeranyl diphosphate synthases expressed in different organs are localized into three subcellular compartments in *Arabidopsis*. Plant Physiol 122: 1045-1056

Orbach MJ, Farrall L, Sweigard JA, Chumley FG, Valent B (2000) A telomeric avirulence gene determines efficacy for the rice blast resistance gene *Pi-ta*. Plant Cell 12: 2019-2032

Panaccione DG, Johnson RD, Wang J, Young CA, Damrongkool P, Scott B, Schardl CL (2001) Elimination of ergovaline from a grass-*Neotyphodium* endophyte symbiosis by genetic modification of the endophyte. Proc Natl Acad Sci USA 98: 12820-12825

Panaccione DG, Pitkin JW, Walton JD, Annis SL (1996) Transposon-like sequences at the TOX2 locus of the plant-pathogenic fungus *Cochliobolus carbonum*. Gene 176: 103-109

Panaccione DG, Scott-Craig JS, Pocard J, Walton JD (1992) A cyclic peptide synthetase gene required for pathogenicity of the fungus *Cochliobolus carbonum* on maize. Proc Natl Acad Sci USA 89: 6590-6594

Panaccione DG, Tapper BA, Lane GA, Davies E, Fraser K (2003) Biochemical outcome of blocking the ergot alkaloid pathway of a grass endophyte. J Agric Food Chem. 51: 6429-6437

Parker EJ, Scott DB (2004) Indole-diterpene biosynthesis in ascomycetous fungi. In Z An, ed, Handbook of Industrial Mycology Vol Chapter 14; Chapter 14. Marcel Dekker, New York, pp 405-426

Payne GA, Nystrom GJ, Bhatnagar D, Cleveland TE, Woloshuk CP (1993) Cloning of the *afl-2* gene involved in aflatoxin biosynthesis from *Aspergillus flavus*. Appl Environ Microbiol 59: 156-162

Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 85: 2444-2448

Pedley KF, Walton JD (2001) Regulation of cyclic peptide biosynthesis in a plant pathogenic fungus by a novel transcription factor. Proc Natl Acad Sci USA 98: 14174-14179

Penn J, Garthwaite I, Christensen MJ, Johnson CM, Towers NR (1993) The importance of paxilline in screening for potentially tremorgenic *Acremonium* isolates. In DE Hume, GCM Latch, HS Easton, eds, Proceedings of the Second International Symposium on Acremonium/Grass Interactions. AgResearch, Grasslands Research Centre, Palmerston North, New Zealand, pp 88-92

Peplow AW, Meek IB, Wiles MC, Phillips TD, Beremand MN (2003a) Tri16 is required for esterfication of postion C-8 during trichothecene mycotoxin production by *Fusarium sporotrichioides*. Appl Environ Microbiol 69: 5935-5940

Peplow AW, Tag AG, Garifullina GF, Beremand MN (2003b) Identification of new genes positively regulated by Tri10 and a regulatory network for trichothecene mycotoxin production. Appl Environ Microbiol 69: 2731-2736

Philipson MN, Christey MC (1986) The relationship of host and endophyte during flowering, seed formation, and germination on *Lolium perenne*. N Z J Bot 24: 125-134

Pirttilä AM, McIntyre LM, Payne GA, Woloshuk CP (2004) Expression profile analysis of wild-type and *fcc1* mutant strains of *Fusarium verticillioides* during fumonisin biosynthesis. Fungal Gen Biol 41: 647-656

Pitkin JW, Nikolskaya A, Ahn J, Walton JD (2000) Reduced virulence caused by meiotic instability of the Tox2 chromosome of the maize pathogen *Cochliobolus carbonum*. Mol Plant Microbe Interact 13: 80-87

Pitkin JW, Panaccione DG, Walton JD (1996) A putative cyclic peptide efflux pump encoded by the *TOXA* gene of the plant-pathogenic fungus *Cochliobolus carbonum*. Microbiology 142: 1557-1565

Popay AJ, Hume DE, Baltus JG, Latch GCM, Tapper BA, Lyons TB, Cooper BM, Pennell CG, Eerens JPJ, Marshall SL (1999) Field performance of perennial ryegrass (*Lolium perenne*) infected with toxin-free fungal endophytes (*Neotyphodium* spp.). In DR Woodfield, C Matthew, eds, Ryegrass endophyte: an essential New Zealand symbiosis. New Zealand Grassland Association, Napier, New Zealand, pp 113-122

Proctor RH, Brown DW, Plattner RD, Desjardins AE (2003) Co-expression of 15 contiguous genes delineates a fumonisin biosynthetic gene cluster in *Gibberella moniliformis*. Fungal Gen Biol 38: 237-249

Proctor RH, Desjardins AE, Plattner RD, Hohn TM (1999) A polyketide synthase gene reuqired for biosynthesis of fumonisin mycotoxins in *Gibberella fujikuroi* mating population A. Fungal Gen Biol 27: 100-112

Proctor RH, Hohn TM, McCormick SP, Desjardins AE (1995) *Tri6* encodes an unusual zinc finger protein involved in regulation of trichothecene biosynthesis in *Fusarium sporotrichioides*. Appl Environ Microbiol 61: 1923-1930

Reddy PV, Lam CK, Belanger FC (1996) Mutualistic fungal endophytes express a proteinase that is homologous to proteases suspected to be important in fungal pathogenicity. Plant Physiol 111: 1209-1218

Reinholz J, Paul VH (2001) Toxin-free *Neotyphodium*-isolates achieved without genetic engineering - a possible strategy to avoid "ryegrass staggers". In VH Paul, PD Dapprich, eds, 4th International Neotyphodium/grass interactions Symposium. Universität Paderborn, Soest, Germany, pp 261-271

Rojas MC, Hedden P, Gaskin P, Tudzynski B (2001) The P450-1 gene of *Gibberella fujikuroi* encodes a multifunctional enzyme in gibberellin biosynthesis. Proc Natl Acad Sci USA 98: 5838-5843

Rojas MC, Urrutia O, Cruz C, Gaskin P, Tudzynski B, Hedden P (2004) Kaurenolides and fujenoic acids are side products of the gibberellin P450-1 monooxygenase in *Gibberella fujikuroi*. Phytochemistry 65: 821-830

Rose MS, Yun S-, Asvarak T, Lu S-, Yoder OC, Turgeon BG (2002) A decarboxylase encoded at the *Cochliobolus heterostrophus* translocation-associated Tox1B locus is required for polyketide (T-toxin) biosynthesis and high virulence on T-cytoplasm maize. Mol Plant Microbe Interact 15: 883-893

Rosewich UL, Kistler HC (2000) Role of horizontal gene transfer in the evolution of fungi. Annu Rev Phytopathol 38: 325-363

Rowan DD, Latch GCM (1994) Utilization of endophyte-infected perennial ryegrasses for increased insect resistance. In CW Bacon, JF White Jr, eds, Biotechnology of Endophytic Fungi of Grasses. CRC Press, Boca Raton, pp 169-183

Schafer WR, Rine J (1992) Protein prenylation: genes, enzymes, targets, and functions. Annu Rev Genet 30: 209-237

Schardl CL, Craven KD (2003) Interspecific hybrisization in plant-associated fungi and oomycetes: a review. Mol Ecol 12: 2861-2873

Schardl CL, Leuchtmann A, Tsai H-, Collett MA, Watt DM, Scott DB (1994) Origin of a fungal symbiont of perennial ryegrass by interspecific hybridization of a mutualist with the ryegrass choke pathogen, *Epichloë typhina*. Genetics 136: 1307-1317

Scott B (2003) The evolution of gene clusters in plant-associated filamentous fungi. NZ BioScience 12: 22-25

Scott-Craig JS, Panaccione DG, Pocard JA, Walton JD (1992) The cyclic peptide synthetase catalyzing HC-toxin production in the filamentous fungus *Cochliobolus carbonum* is encoded by a 15.7-kilobase open reading frame. J Biol Chem 267: 26044-26049

Selker EU (1990) Premeiotic instability of repeated sequences in *Neurospora crassa*. Annu Rev Genet 24: 579-613

Seo J-, Proctor RH, Plattner RD (2001) Characterization of Four Clustered and Coregulated Genes Associated with Fumonisin Biosynthesis in *Fusarium verticillioides*. Fungal Gen Biol 34: 155-165

Seya H, Nozawa K, Udagawa S, Nakajima S, Kawai K (1986) Studies on fungal products. IX. Dethiosecoemestrin, a new metabolite related to emestrin, from *Emericella striata*. Chem Pharm Bull 34: 2411-2416

Siegel MR, Bush LP (1997) Toxin production in grass/endophyte associations. In Carroll, Tudzynski, eds, The Mycota V part A. Springer-Verlag, Berlin, pp 185-207

Siegel MR, Latch GCM, Bush LP, Fannin FF, Rowan DD, Tapper BA, Bacon CW, Johnson MC (1990) Fungal endophyte-infected grasses: alkaloid accumulation and aphid response. J Chem Ecol 16: 3301-3315

Skory CD, Chang P-, Linz JE (1993) Regulated expression of the *nor-1* and *ver-1* genes associated with aflatoxin biosynthesis. Appl Environ Microbiol 59: 1642-1646

Snoeijers SS, Perez-Garcia A, Goosen T, De Wit PJGM (2003) Promoter analysis of the avirulence gene Avr9 of the fungal tomato pathogen *Cladosporium fulvum* in the model filamentous fungus *Aspergillus nidulans*. Curr Genet 43: 96-102

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98: 503-517

Spiering MJ, Davies E, Tapper BA, Schmid J, Lane GA (2002a) Simplified extraction of ergovaline and peramine for analysis of tissue distribution in endophyte-infected grass tillers. J Agric Food Chem 50: 5856-5862

Spiering MJ, Lane GA, Christensen MJ, 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 MJ, Moon CD, Wilkinson HH, Schardl CL (2005b) Gene clusters for insecticidal loline alkaloids in the grass-endophytic fungus *Neotyphodium uncinatum*. Genetics 196: 1403-1414

Spiering MJ, Wilkinson HH, Blankenship JD, Schardl CL (2002b) Expressed sequence tags and genes associated with loline alkaloid expression by the fungal endophyte *Neotyphodium uncinatum*. Fungal Gen Biol 36: 242-254

Sweigard JA, Carroll AM, Kang S, Farrall L, Chumley FG, Valent B (1995) Identification, cloning, and characterization of *PWL2*, a gene for host species specificity in the rice blast fungus. Plant Cell 7: 1221-1233

Tag A, Hicks J, Garifullina G, Ake C,Jr, Phillips TD, Beremand M, Keller N (2000) G-protein signalling mediates differential production of toxic secondary metabolites. Mol Microbiol 38: 658-665

Tag AG, Garifullina GF, Peplow AW, Ake C,Jr, Phillips TD, Hohn TM, Beremand MN (2001) A novel regulatory gene, *Tri10*, controls trichothecene toxin production and gene expression. Appl Environ Microbiol 67: 5294-5302

Takeshita N, Ohta A, Horiuchi H (2002) *csmA*, a gene encoding a class V chitin synthase with a myosin motor-like domain of *Aspergillus nidulans*, is translated as a single polypeptide and regulated in response to osmotic conditions. Biochem Biophys Res Commun 298: 103-109

Tan YY, Spiering MJ, Scott V, Lane GA, Christensen MJ, Schmid J (2001) *In planta* regulation of extension of an endophytic fungus and maintainance of high metabolic rates in its mycelium in the absences of apical extension. Appl Environ Microbiol 67: 5377-5383

Tanaka A, Shiotani H, Yamamoto M, Tsuge T (1999) Insertional mutagenesis and cloning of the genes required for biosynthesis of the host-specific AK-toxin in the Japanese pear pathotype of *Alternaria alternata*. Mol Plant Microbe Interact 12: 691-702

Tanaka A, Tsuge T (2001) Reporter gene analysis of AKT3-1 and AKT3-2 expression during conidial germination of the Japanese pear pathotype of *Alternaria alternata.* J Gen Plant Path 67: 15-22

Tanaka A, Tsuge T (2000) Structural and functional complexity of the genomic region controlling AK-toxin biosynthesis and pathogenicity in the Japanese pear pathotype of *Alternaria alternata.* Mol Plant Microbe Interact 13: 975-986

Tapper BA, Latch GCM (1999) Selection against toxin production in endophyte-infected perennial ryegrass. Eds Woodfield DR, Matthew C In: Procedings: Ryegrass: an essential New Zealand symbiosis. Grasslands research and practice series No 7. 107-111

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nculeic Acids Res 22:4673-4680

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876-4882

Tilburn J, Sarkar S, Widdick DA, Espeso EA, Orejas M, Mungroo J, Peñalva MA, Arst HN (1995) The *Aspergillus* PacC zinc finger transcription factor mediates regulation of both acid- and alkaline-expressed genes by ambient pH. EMBO J 14: 779-790

Todd RB, Andrianopoulos A (1997) Evolution of a fungal regulatory gene family: The Zn(II)2Cys6 binuclear cluster DNA binding motif. Fungal Gen Biol 21: 388-405

Toyomasu T, Nakaminami K, Toshima H, Mie T, Watanabe K, Ito H, Matsui H, Mitsuhashi W, Sassa T, Oikawa H (2004) Cloning of a gene cluster responsible for the biosynthesis of diterpene aphidicolin, a specific inhibitor of DNA polymerase alpha. Biosci Biotechnol Biochem 68: 146-152

Trail F, Chang PK, Cary J, Linz JE (1994) Structural and functional analysis of the *nor-1* gene involved in the biosynthesis of aflatoxins by *Aspergillus parasiticus.* Appl Environ Microbiol 60: 4078-4085

Trapp SC, Hohn TM, McCormick S, Jarvis BB (1998) Characterization of the gene cluster for biosynthesis of macrocyclic trichothecenes in *Myrothecium roridum*. Mol Gen Genetics 257: 421-432

Tsai H-, Liu J-, Staben C, Christensen MJ, Latch GCM, Siegel MR, Schardl CL (1994) Evolutionary diversification of fungal endophytes of tall fescue grass by hybridization with *Epichloë* species. Proc Natl Acad Sci USA 91: 2542-2546

Tsai H-, Wang H, Gebler JC, Poulter CD, Schardl CL (1995) The *Claviceps purpurea* gene encoding dimethylallyltryptophan synthase, the committed step for ergot alkaloid biosynthesis. Biochem Biophys Res Commun 216: 119-125

Tudzynski B, Hedden P, Carrera E, Gaskin P (2001) The P450-4 gene of *Gibberella fujikuroi* encodes *ent*-kaurene oxidase in the gibberellin biosynthesis pathway. Appl Environ Microbiol 67: 3514-3522

Tudzynski B, Hölter K (1998) Gibberellin biosynthetic pathway in *Gibberella fujikuroi*: evidence for a gene cluster. Fungal Gen Biol 25: 157-170

Tudzynski B, Honman B, Feng B, Marzluf GA (1999) Isolation, characterisation and disruption of the areA nitrogen regulatory gene of *Gibberella fujikuroi*. Mol Gen Genet 261: 106-114

Tudzynski B, Mihlan M, Rojas MC, Linnemannstoens P, Gaskin P, Hedden P (2003) Characterization of the final two genes of the gibberellin biosynthesis gene cluster of *Gibberella fujikuroi*: *des* and P450-3 encode GA4 desaturase and the 13-hydroxylase respectively. J Biol Chem 278: 28635-28643

Tudzynski B, Rojas MC, Gaskin P, Hedden P (2002) The gibberellin 20-oxidase of *Gibberella fujikuroi* is a multifunctional monooxygenase. J Biol Chem 277: 21246-21253

Tudzynski P, Correia T, Keller U (2001) Biotechnology and genetics of ergot alkaloids. Appl Microbiol Biotechnol 57: 593-605

Tudzynski P, Hölter K, Correia T, Arntz C, Grammel N, Keller U (1999) Evidence for an ergot alkaloid gene cluster in *Claviceps purpurea*. Mol Gen Genet 261: 133-141

van den Broek P, Pittet A, Hajjaj H (2001) Aflatoxin genes and the aflatoxigenic potential of koji moulds. Appl Environ Microbiol 57: 192-199

van Helden J (2003) Regulatory sequence analysis tools. 2003 31: 3593-3596

van Helden J, Andre B, Collado-Vides J (2000a) A web site for the computational analysis of yeast regulatory sequences. Yeast 16: 177-187

van Helden J, Andre B, Collado-Vides J (1998) Extracting regulatory sites from the upstram region of yeast genes by computational analysis of oligonucleotide frequencies. J Mol Biol 281: 827-842

van Helden J, Rios AF, Collado-Vides J (2000b) Discovering regulatory elements in non-coding sequences by analysis of spaced dyads. Nucleic Acids Res 28: 1808-1818

van Zijll de Jong,E., Guthridge KM, Spangenberg GC, Forster JW (2003) Development and characterization of EST-derived simple sequence repeat (SSR) markers for pasture grass endophytes. Genome 46: 277-290

Vollmer SJ, Yanofsky C (1986) Efficient cloning of genes of *Neurospora crassa*. Proc Natl Acad Sci USA 83: 4869-4873

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

Walton JD (2000) Horizontal gene transfer and the evolution of secondary metabolite gene clusters in fungi: an hypothesis. Fungal Gen Biol 30: 167-171

Wang J, Machado C, Panaccione DG, Tsai H, Schardl CL (2004) The determinant step in ergot alkaloid biosynthesis by an endophyte of perennial ryegrass. Fungal Gen Biol 41: 189-198

Watson AJ, Fuller LJ, Jeenes DJ, Archer DB (1999) Homologs of aflatoxin biosynthesis genes and sequence of *aflR* in *Aspergillus oryzae* and *Aspergillus sojae*. Appl Environ Microbiol 65: 307-310

Weedon CM, Mantle PG (1987) Paxilline biosynthesis by *Acremonium loliae*; a step towards defining the origin of lolitrem neurotoxins. Phytochemistry 26: 969-971

Wilkinson HH, Siegel MR, Blankenship JD, Mallory AC, Bush LP, Schardl CL (2000) Contribution of fungal loline alkaloids to protection from aphids in a grass-endophyte mutualism. Mol Plant Microbe Interact 13: 1027-1033

Woloshuk CP, Fount KR, Brewer JF, Bhatnagar D, Cleveland TE, Payne GA (1994) Molecular characterisation of *aflR*, a regulatory locus for aflatoxin biosynthesis. Appl Environ Microbiol 60: 2408-2414

Woloshuk CP, Prieto R (1998) Genetic organization and function of the aflatoxin B1 biosynthetic genes. FEMS Microbiol Lett 160: 169-176

Yang G, Rose MS, Turgeon BG, Yoder OC (1996a) A polyketide synthase is required for fungal virulence and production of the polyketide T-toxin. Plant Cell 8: 2139-2150

Yang T, Cheng L, Kain SR (1996b) Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein. Nucleic Acids Res 24: 4592-4593

Yelton MM, Hamer JE, Timberlake WE (1984) Transformation of *Aspergillus nidulans* by using a *trpC* plasmid. Proc Natl Acad Sci USA 81: 1470-1474

Yoder OC (1988) *Cochliobolus heterostrophus*, cause of southern corn leaf blight. Advan Plant Path 6: 93-112

Young C, Itoh Y, Johnson R, Garthwaite I, Miles CO, Munday-Finch SC, Scott B (1998) Paxilline-negative mutants of *Penicillium paxilli* generated by heterologous and homologous plasmid integration. Curr Genet 33: 368-377

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

Yu J, Cary JW, Bhatnagar D, Cleveland TE, Keller NP, Chu FS (1993) Cloning and characterisation of a cDNA from *Aspergillus parasiticus* encoding an O-methyltransferase involved in aflatoxin biosynthesis. Appl Environ Microbiol 59: 3564-3571

Yu J, Chang P, Bhatnagar D, Cleveland TE (2000) Cloning of a sugar utilization gene cluster in *Aspergillus parasiticus*. Biochim Biophys Acta 1493: 211-214

Yu J, Chang P-, Cary JW, Bhatnagar D, Cleveland TE (1997) *avnA*, a gene encoding a cytochrome P-450 monooxygenase, is involved in the conversion of averantin to averufin in aflatoxin biosynthesis in *Aspergillus parasiticus*. Appl Environ Microbiol 63: 1349-1356

Yu J, Chang P, Cary JW, Wright M, Bhatnagar D, Cleveland TE, Payne GA, Linz JE (1995) Comparitive mapping of aflatoxin pathway gene clusters in *Aspergillus parasiticus* and *Aspergillus flavus*. Appl Environ Microbiol 61: 2365-2371

Yu J, Chang P, Ehrlich KC, Cary JW, Bhatnagar D, Cleveland TE, Payne GA, Linz JE, Woloshuk CP, Bennett JW (2004) Clustered pathway genes in aflatoxin biosynthesis. Appl Environ Microbiol 70: 1253-1262

Yu J, Chang P, Ehrlich KC, Cary JW, Montalbano B, Dyer JM, Bhatnagar D, Cleveland TE (1998) Characterisation of the critical amino acids of an *Aspergillus parasiticus* cytochrome P-450 monooxygenase encoded by ordA that is involved in the biosynthesis of aflatoxins B1,G1,B2 and G2. Appl Environ Microbiol 64: 4834-4841

Yu J, Butchko RAE, Fernandes M, Keller NP, Leonard TJ, Adams TH (1996) Conservation of structure and function of the aflatoxin regulatory gene *aflR* from *Aspergillus nidulans* and *A. flavus*. Curr Genet 29: 549-555

Yu J-, Leonard TJ (1995) Sterigmatocystin biosynthesis in *Aspergillus nidulans* requires a novel type I polyketide synthase. J Bacteriol 177: 4792-4800

Zhang S, Monahan BJ, Tkacz JS, Scott B (2004) Molecular cloning and genetic analysis of an indole-diterpene gene cluster from *Aspergillus flavus*. Appl Environ Microbiol 70: 6875-6883

Zhang X (2004) Functional analysis of a thiamine biosynthetic gene in the interaction of *Epichloë typhina* with perennial ryegrass. PhD thesis. Massey University.

Zolan ME (1995) Chromosome-length polymorphism in fungi. Microbiol Rev 1995: 686-698

Molecular Genetics and Genomics

DOI 10.1007/s00438-005-1130-0

Original Paper

Molecular cloning and genetic analysis of a symbiosis-expressed gene cluster for lolitrem biosynthesis from a mutualistic endophyte of perennial ryegrass

C. A. Young • M. K. Bryant • M. J. Christensen • B. A. Tapper · G. T. Bryan · B. Scott (✉)

C. A. Young · M. K. Bryant · B. Scott
Centre for Functional Genomics, Institute of Molecular BioSciences, College of Sciences, Massey University, Private Bag 11 222, Palmerston North, New Zealand

M. J. Christensen · B. A. Tapper · G. T. Bryan
AgResearch Grasslands Research Centre, Tennent Drive, Private Bag 11 008, Palmerston North, New Zealand

✉ B. Scott
Phone: +64-6-3505168
Fax: +64-6-3505688
E-mail: d.b.scott@massey.ac.nz

Received: 1 November 2004 / **Accepted:** 11 February 2005

Abstract Lolitrems are potent tremorgenic mycotoxins that are synthesised by clavicipitaceous fungal endophytes of the *Epichloë/Neotyphodium* group in association with grasses. These indole-diterpenes confer major ecological benefits on the grass–endophyte symbiotum. A molecular signature for diterpene biosynthesis is the presence of two geranylgeranyl diphosphate (GGPP) synthases. Using degenerate primers for conserved domains of fungal GGPP synthases, we cloned two such genes, *ltmG* and *ggsA*, from *Neotyphodium lolii*. Adjacent to *ltmG* are two genes, *ltmM* and *ltmK*, that are predicted to encode an FAD-dependent monooxygenase and a cytochrome P450 monooxygenase, respectively. The cluster of *ltm* genes is flanked by AT-rich retrotransposon DNA that appears to have undergone extensive repeat induced point (RIP) mutation. *Epichloë festucae*, the sexual ancestor of *N. lolii*, contains an identical *ltm* gene cluster, but lacks the retrotransposon "platform" on the right flank. Associations established between perennial ryegrass and an *E.*

festucae mutant deleted for *ltmM* lack detectable levels of lolitrems. A wild-type copy of *ltmM* complemented this phenotype, as did *paxM* from *Penicillium paxilli*. Northern hybridization and RT-PCR analysis showed that all three genes are weakly expressed in culture but strongly induced in planta. The relative endophyte biomass in these associations was estimated by real-time PCR to be between 0.3 and 1.9%. Taking this difference into account. the steady-state levels of the *ltm* transcripts are about 100-fold greater than the levels of the endogenous ryegrass β-tubulin (β -*Tub1*) and actin (*Act1*) RNAs. Based on these results we propose that *ltmG*, *ltmM* and *ltmK* are members of a set of genes required for lolitrem biosynthesis in *E. festucae* and *N. lolii*.

Electronic Supplementary Material Supplementary material is available for this article at http://dx.doi.org/10.1007/s00438-005-1130-0

Keywords *Neotyphodium lolii* · *Epichloë festucae* · Lolitrem B · Endophyte · Retrotransposons

Communicated by P.Punt

Introduction

Lolitrems are mycotoxins that are potent tremorgens in mammals, and are synthesised by clavicipitaceous fungal endophytes of the genera *Epichloë*/ *Neotyphodium* in association with their grass hosts (Fletcher and Harvey 1981; Gallagher et al. 1981, 1984; Munday-Finch et al. 1995). The most abundant of this group of metabolites is lolitrem B (Fig. 1). Lolitrems are members of a large, structurally diverse group of secondary metabolites known as indole–diterpenes that includes paspaline (Cole et al. 1977) and paxilline (Springer et al. 1975). These metabolites have a common structural core comprised of a cyclic diterpene skeleton derived from geranylgeranyl diphosphate (GGPP) and an indole moiety derived from anthranilic acid. Different patterns of prenylation, hydroxylation, epoxidation and acetylation, and differences in ring stereochemistry around the basic indole–diterpene ring structure, are responsible for this structural diversity (Parker and Scott 2004).

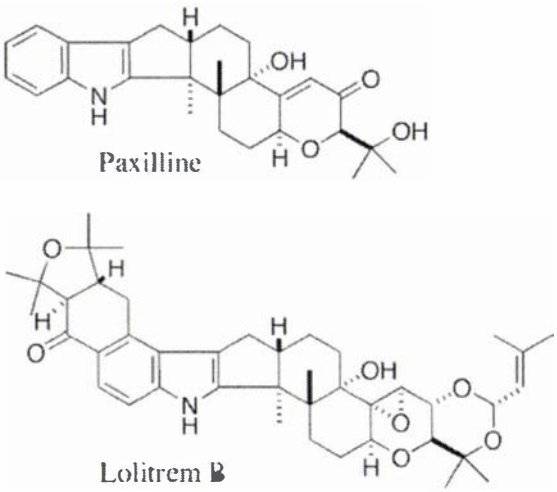


Fig. 1 Structures of paxilline and lolitrem B

These metabolites are proposed to confer major ecological benefits on the grass–endophyte symbiotum by protecting it from mammalian and insect herbivory (Clay 1990; Clay and Schardl 2002). These benefits result in a strong selective pressure for maintenance of the fungal biosynthetic genes (Schardl 1996). Lolitrems are relatively abundant in leaf sheath tissue and seeds of perennial ryegrass (*Lolium perenne*) and some fescue grasses (*Festuca* spp.) containing *Epichloë festucae*, its asexual derivative, *Neotyphodium lolii*, or inter-specific hybrids of these and other *Epichloë*/*Neotyphodium* species (Siegel et al. 1990; Christensen et al. 1993; Bush et al. 1997). There are two reports of indole–diterpene biosynthesis in axenic cultures of *Neotyphodium* spp., which confirm that these compounds are fungal metabolites (Penn et al. 1993; Reinholz and Paul 2001). The preferential synthesis of these metabolites in planta, however, suggests that the genes for lolitrem biosynthesis (designated *ltm* genes) are symbiotically regulated.

From an agricultural perspective, fungal synthesis of indole–diterpenes in forage grasses poses a significant health problem for grazing livestock (Fletcher and Harvey 1981). Lolitrem B is proposed to be the causative agent of "ryegrass staggers syndrome" (Gallagher et al. 1982, 1984). This is a condition in which animals grazing on endophyte-infected pastures develop ataxia, tremors and sensitivity to external stimuli. The reaction to lolitrem is long-lasting, but is completely reversible (Smith et al. 1997; McLeay et al. 1999). The mechanism by which lolitrem B and related indole–diterpenes cause tremors in mammals is not well defined, but biochemical and clinical studies indicate that these effects are due in part to actions on receptors and interference with neurotransmitter release in the central and peripheral nervous systems (Selala et al. 1991).

While the derivation of these compounds from GGPP and indole has been clearly established, very little is known about the pathways for their biosynthesis (Acklin et al. 1977; de Jesus et al. 1983; Byrne et al. 2002). Biosynthetic schemes based on the chemical identification of likely intermediates have been proposed, but until recently none of the proposed steps had been validated by biochemical or genetic studies (Mantle and Weedon 1994; Munday-Finch et al. 1996).

Recently, a cluster of genes for paxilline (Fig. 1) biosynthesis was cloned from *Penicillium paxilli* (Young et al. 2001). Key genes in this cluster include those for a GGPP synthase (*paxG*), an FAD-dependent monooxygenase (*paxM*), a prenyl transferase (*paxC*) and two cytochrome P450 monooxygenases (*paxP* and *paxQ*). Deletion of *paxG*, *paxM* or *paxC* results in mutants that are defective in paxilline biosynthesis (Young et al. 2001; B. Scott, L.K. McMillan, J.W. Astin, S. Saikia, C.A. Young, A. Bryant and E.J. Parker, unpublished results). PaxG is proposed to catalyse the determinant step in paxilline biosynthesis. PaxM and PaxC are proposed to catalyse the addition of indole-3-glycerol phosphate to GGPP and subsequent cyclisation to form the first stable indole-diterpene, possibly paspaline (Parker and Scott 2004). The Δ*paxP* and Δ*paxQ* mutants accumulate paspaline and 13-desoxypaxilline, respectively, suggesting that these compounds are the substrates for the corresponding enzymes (McMillan et al. 2003). Thus, at least five genes are required for paxilline biosynthesis in *P. paxilli*.

Three other fungal gene clusters for cyclic diterpene biosynthesis have been reported, one for gibberellin biosynthesis in *Fusarium fujikuroi* (teleomorph *Gibberella fujikuroi*) (Tudzynski and Hölter 1998), one for aphidicolin biosynthesis in *Phoma betae* (Toyomasu et al. 2004) and one for aflatrem biosynthesis in *Aspergillus flavus* (Zhang et al. 2004). Like the paxilline biosynthetic cluster in *P. paxilli*, the gibberellin, aphidicolin and aflatrem gene clusters contain a gene for a GGPP synthase (Tudzynski and Hölter 1998; Toyomasu et al. 2004; Zhang et al. 2004). Furthermore, *P. paxilli*, *F. fujikuroi* and *A. flavus* each contain a second GGPP synthase gene, suggesting that the presence of two GGPP synthase genes may be a molecular signature for diterpene biosynthesis. Given that genes for secondary metabolite biosynthesis in fungi are generally organised in clusters (Keller and Hohn 1997), molecular cloning of GGPP synthase genes, in combination with chromosome walking, provides a rapid strategy for cloning new indole–diterpene gene clusters.

The objectives of this present study were: (1) to isolate GGPP synthase genes from *N. lolii* and *E. festucae* by degenerate PCR, (2) to determine if any of the GGPP synthase genes is closely linked to homologues of the *P. paxilli pax* genes by chromosome walking, (3) to determine whether these genes are preferentially expressed in planta, and (4) to confirm that at least one of the genes

is necessary for lolitrem biosynthesis by deletion and complementation analysis. Our working hypothesis was that *N. lolii* and *E. festucae* contain two GGPP synthase genes, one of which is associated with a cluster of genes for lolitrem biosynthesis that are preferentially expressed in planta.

Materials and methods

Bacterial strains and plasmids

The *Escherichia coli* strains XL1-Blue (Bullock et al. 1987) and KW251 (Promega) were grown on LB agar plates and, where necessary, the medium was supplemented with ampicillin (100 μg/ml).

Fungal strains and growth conditions

Cultures of *N. lolii* strain Lp19 (Christensen et al. 1993), *E. festucae* strain Fl1 (ex cultivar SR3000) and derivatives (this study), *E. typhina* strain E8 (Schardl et al. 1994), and the *P. paxilli* strains PN2013 and LM662 (Young et al. 2001) were grown and maintained as previously described (Moon et al. 1999, 2000; Young et al. 2001; McMillan et al. 2003). Cultures of Lp19 and Fl1 used for RT-PCR analysis were grown either on potato dextrose (PD) medium or first on PD, then washed and transferred to a defined medium containing CD (Czapek-Dox) salts plus a nitrogen and/or carbon source as described below. CD salts comprised 4.4 mM K_2HPO_4, 2 mM $MgSO_4 \cdot 7H_2O$, 6.7 mM KCl and 36 μM $FeSO_4$.

Plant growth conditions and inoculations

Inoculation of endophyte-free seedlings of perennial ryegrass (cv. Nui) was carried out by the method of Latch and Christensen (1985) using mycelium from a 7-day-old culture. Plants were tested for endophyte infection by tissue-print immunoblotting (Gwinn et al. 1991) using polyclonal rabbit antibodies raised against homogenised mycelium of *N. lolii* strain Lp5 (Christensen et al. 1993), in conjunction with a goat anti-rabbit secondary antibody linked to an alkaline phosphatase conjugate (Sigma). Confirmation that plants were infected with endophtye was obtained by light microscopic examination of epidermal strips from the outermost leaf sheath after staining with aniline blue.

Molecular biological techniques

Genomic DNA from *N. lolii*, *E. festucae*, *E. typhina* and *P. paxilli* was isolated from freeze-dried mycelium using previously described methods (Yoder 1988; Byrd et al. 1990; Möller et al. 1992). Plasmid DNA was isolated and purified using either a BioRad Quantum Prep plasmid miniprep kit (BioRad Laboratories) or a Qiagen plasmid mini kit. PCR products amplified with *Taq* DNA polymerase (Roche Diagnostics) were routinely cloned into pGEM-T, pGEM-T easy (Promega) or pUC118, and transformed into *E. coli* XL1-Blue. Genomic digests were transferred to positively charged nylon membranes (Roche) by capillary transfer (Southern 1975) and fixed by UV crosslinking (120,000 $\mu J/cm^2$) in a Ultraviolet crosslinker Cex-800 (Ultra-Lum Inc.). Filters were probed with [$\alpha-^{32}$ P]dCTP (3,000 Ci/mmol; Amersham) labelled probes. Templates were purified using a PCR purification kit (Qiagen). DNA was labelled by randomly primed synthesis using a High-Prime kit (Roche). Labelled probes were purified using ProbeQuant columns (Pharmacia). Hybridisations were carried out as described by Young et al. (1998).

Construction and screening of a Lp19 genomic library

Genomic DNA isolated from *N. lolii* strain Lp19 was partially digested with *Mbo*I using the method described by Frischauf et al. (1983) to generate the maximum yield of fragments in the size range of 9–23 kb. This DNA was partially end-filled with dATP and dGTP using the Klenow fragment of DNA polymerase I (Roche) to generate 5′-GA protruding termini. Aliquots of this DNA were ligated in a 1:2 molar ratio of vector to insert with 1 μg of λGEM-12 *Xho*I half-site arms (Promega). The ligated mixture was packaged using the Packagene system (Promega) according to the manufacturer's instructions. The resulting phages were transfected into *E. coli* strain KW251 and 2.2×10^5 plaques were obtained. These were washed off the plates and stored in 7% DMSO at −20°C. The library was screened by plaque hybridisation using standard methods (Sambrook et al. 1989).

Polymerase chain reaction

Standard PCR amplifications of genomic, plasmid and bacteriophage DNA templates were carried out in 25-μl reactions containing 10 mM Tris–HCl, 1.5 mM $MgCl_2$, and 50 mM KCl (pH 8.3), each dNTP at 50 μM and each primer at 200 nM, 0.5 U of *Taq* DNA polymerase (Roche) and 5 ng of genomic DNA. The thermocycle conditions used were: one cycle at 94°C for 2 min; 30 cycles at 94°C for 15 s, 60°C for 30 s, and 72°C for 1 min (per kb), followed by a final step at

72°C for 5 min. Reaction and cycle conditions for PCRs using degenerate primers were the same as described above, except that the final primer concentration in the reaction mixture was 800 nM and the cycle annealing temperature was reduced to 45°C. The nucleotide sequences of the degenerate primers ggpps27, ggpps28 and ggpps29 (Table S1) were based on conserved domains identified from an alignment of fungal GGPP synthase polypeptide sequences (Zhang et al. 2004).

Inverse PCR was carried out under standard reaction and cycle conditions using 50 ng of self-ligated genomic DNA as template, with primers CYLp19-18 and CYLp19-16 (Table S1). For the analysis of the *ggsA* sequences in *E. typhina* and *E. festucae* an *Sst*I digest was used with the amplified product of inverse PCR as probe.

RNA isolation, cDNA synthesis and expression analysis

RNA was isolated from frozen fungal mycelium or pseudostem tissue of perennial ryegrass, using the TRIzol reagent (Invitrogen). DNA was removed by incubating ~30 μg aliquots of total RNA with 30 U of Dnase1 (Roche) for 30 min at 37°C, followed by 10 min at 75°C. mRNA was isolated by oligo(dT) affinity-chromatography using a Sigma GenElute mRNA miniprep kit. Samples of denatured mRNA (~50 ng) were reverse-transcribed at 42°C for 45 min in a reaction volume of 20 μl containing 1× reaction buffer (Roche), 10 mM DTT, each dNTP at 50 μM, 1.8 nmol of random primers (Roche), 8 U of RNase inhibitor (Roche) and 50 U of Expand Reverse Transcriptase (Roche). Gene-specific amplification of cDNA dilutions (1/10, 1/100 and 1/1,000) was carried out in a 25-μl volume containing 1× *Taq* polymerase buffer (Roche), each dNTP at 50 μM, each primer at 200 nM and 0.5 U of *Taq* polymerase (Roche). The thermocycle conditions used were: one cycle at 94°C for 2 min, followed by 35 cycles at 94°C for 15 s, 60°C for 30 s and 72°C for 45 s; followed by a final incubation at 72°C for 10 min. For intron analysis the cDNA was amplified with the following primer pairs (Table S1): for *ltmG*, lol79 and lol1 (introns 1and 2); for *ltmM*, lol7 and lol35 (intron 1), and lol14 and lol28 (introns 2 and 3); for *ltmK*, lol29 and lol15 (introns 1–5), and lol43 and lol63 (introns 6 and 7). For expression analysis the cDNA was amplified with the following primer pairs: for *ltmG*, lol3 and lol1; for *ltmM*, lol14 and lol28; for *ltmK*, lol32 and lol15; for the endophyte *tubB* gene, T1.1 and T1.2; and for the β-*Tub1* gene of perennial ryegrass, PRG3 and PRG4.

To analyse *ggsA* transcripts from Lp19 and Fl1, one-step reverse transcriptase-PCR was carried out using Superscript II with Platinum *Taq* (Invitrogen) and total mycelium RNA (1 μg) under the reaction conditions recommended by the manufacturer. The cycle conditions used were: one cycle at 50°C for 20 min; one cycle at 94°C for 2 min; 35 cycles at 94°C for 15 s, 55°C for 30 s,

and 72°C for 1 min; with a final incubation at 72°C for 10 min. The primer pairs used were CYLp19-22 and CYLp19-18, and CYLp19-16 and CYLp19-1 (Table S1).

For Northern analysis total RNA was fractionated by electrophoresis for 4–5 h at 120 V on a formaldehyde (6.2%)-agarose (1.2%) gel containing 1×MOPS buffer. The denatured RNA was transferred to a positively charged nylon membrane (Roche) by capillary transfer and the RNA was fixed by UV crosslinking. Filters were hybridised with cDNA probes labelled with [α-32P]dCTP (3,000 Ci/mmol; Amersham). Hybridisations were carried out at 65°C overnight in 5×SSPE containing 5′ Denhardt's solution and 0.5% (w/v) SDS. Filters were washed once in 1×SSC–0.1% (w/v) SDS at 68°C for 20 min, then three times in 0.2×SSC–0.1% (w/v) SDS at 68°C. The primer combinations (Table S1) used to generate the cDNA (*ltmG*, *ltmM* and *ltmK*) and genomic (*tubB* and *Act1*) probes were amplified with the following primer pairs: for *ltmG*, lol79 and lol1; for *ltmM*, lol7 and lol28; for *ltmK*, lol32 and lol8; for Lp19 *tubB*, T1.1 and T1.2; and for perennial ryegrass *Act1*, actinF and actinR.

Preparation of deletion and complementation constructs

Plasmid pPN1688 was constructed by ligating a 1.4-kb *Bam*HI fragment containing PtrpC*hph* into pUC118. The *Bam*HI fragment was prepared by digesting a PCR product amplified with the primer set CY4 and CY5 (Table S1) from plasmid pCB1004 (Carroll et al. 1994).

Plasmid pCY39 (the *ltmM* gene replacement vector) was constructed by sequentially ligating into pPN1688 a 2.7-kb *Pst*I fragment 3′ of *ltmM*, and a 2.7-kb *Kpn*I fragment 5′ of *ltmM*. The two fragments were prepared by digesting PCR products amplified from FII genomic DNA with primer set lol17 and lol18 (3,145 bp), and primer set lol48 and lol49 (2,746 bp), respectively. The primer sets contained mismatches relative to the plasmid or genomic DNA in order to introduce the appropriate enzyme recognition site (Table S1).

PCR amplifications of the 2.7-kb *Pst*I and 2.7-kb *Kpn*I fragments used to construct pCY39 were carried out in a 25-µl reaction volume containing 10 ng of FII genomic DNA, 2.5 µl of 10× reaction buffer (Roche), each primer at 200 nM, each dNTP at 125 µM, and 0.875 U of Expand Hi Fidelity polymerase (Roche). The thermocycle conditions used were: one cycle at 94°C for 2 min; 30 cycles at 94°C for 15 s, 50°C for 30 s, and 68°C for 3.5 min (with a 10-s incremental increase per cycle after cycle 10), followed by incubation at 68°C for 7 min.

The 1.4-kb *Bam*HI fragment was amplified with Expand Hi Fidelity polymerase (Roche) as described above, except that each dNTP was present at a concentration of 50 µM in the reaction

mixture and the thermocycle conditions were: one cycle at 94°C for 2 min; 30 cycles at 94°C for 15 s, 56°C for 30 s and 72°C for 1.75 min; then 72°C for 5 min.

The linear product of pCY39 used for transformation was amplified using Expand Long Template polymerase (Roche) and approximately 1 ng of pCY39 as template as described above, except that each dNTP was added at 350 μM and each primer (M13F and M13R) at 150 nM, and the thermocycle conditions were: one cycle at 93°C for 2 min; 30 cycles at 93°C for 10 s, 55°C for 30 s, 68°C for 6 min (with a 20-s incremental increase per cycle after cycle 12); then 68°C for 10 min.

Plasmid pCY40 was prepared by ligating a 7-kb *Xho*I fragment of λCY218 (see below) into *Xho*I digested pII99 (Namiki et al. 2001). Plasmid pCY41 was prepared by ligating a 3.7 kb *Bgl*II fragment from λCY46 (Young et al. 2001) into the *Bgl*II site of pII99.

Transformation of *E. festucae* and molecular analysis of transformants

Protoplasts of *E. festucae* were prepared as previously described (Young et al. 1998), except that 10 mg/ml Glucanex (Chemcolour) was used to digest the cell walls. Protoplasts were transformed with 5 μg of either linear PCR products (pCY39) or circular plasmids (pCY40 and pCY41) using the method of Vollmer and Yanofsky (1986) as modified by Itoh et al. (1994). Transformants were selected on RG medium containing either hygromycin (150 μg/ml) for replacement constructs, or geneticin (200 μg/ml) for complementation constructs. To obtain clonal isolates the resulting transformants were purified by sub-culturing a 2-mm^2 fragment of mycelium from the edge of a colony onto PD medium containing either hygromycin (150 μg/ml) or geneticin (200 μg/ml). This process was repeated three times. This method was used as *E. festucae* do not produce abundant conidia and it was previously shown that this endophyte species contains a single nucleus per apical compartment (Schmid et al. 2000). Even though mycelia of such monokaryotic fungi can still form heterokaryons, each nucleus type will give rise to a homokaryotic sector within the mycelium (Trinci 1978).

Real-time quantitative PCR determination of endophyte biomass

DNA isolated from perennial ryegrass (Möller et al. 1992) and *E. festucae* FII (Byrd et al. 1990) was resuspended to a concentration of 300 μg/μl, boiled for 5 min, quantified using a DyNA Quant

200 fluorometer (Hoefer) and diluted to 5 ng/μl. To establish a standard curve for endophyte-free perennial ryegrass, samples were prepared by diluting boiled G1138 DNA to a concentration range of 0.01–10 ng/μl. Samples for the endophyte standard curve were prepared by diluting boiled F11 DNA to a concentration range of 0.0015–1.5 ng/μl in a 5 ng/μl solution of G1138 DNA. Amplification of the endophyte and plant amplicons was performed on a Lightcycler (Roche) using a Lightcycler FastStart DNA MasterPLUS SYBR green I system (Roche) with the primer pairs endo1 and endo2 (206 bp) and PRG12 and PRG13 (168 bp), designed to amplify F11 ggsA and perennial ryegrass β-*Tub1* sequences, respectively. Amplifications were carried out in 10-μl reactions containing 2 μl of Lightcycler FastStart DNA Master^{PLUS} SYBR green I system (Roche), each primer at 62.5 nM, and 2 μl of sample DNA. The thermocycle conditions used were: one cycle at 95°C for 10 min; 35 cycles at 95°C for 10 s, 60°C for 5 s and 72°C for 9 or 10 s for perennial ryegrass or endophyte amplifications, respectively, using a single fluorescence acquisition. The melting curve was generated by holding the sample at 65°C for 30 s, followed by ramping up to 95°C at a rate of 0.2°C/s using continuous fluorescence acquisition. Each sample was analysed in duplicate for each primer combination. Quantification was performed using the "fit points" option with the noise bands set to 0.505. The melting curves were used to determine whether there was specific amplification of each product and this was confirmed by electrophoresis of the reaction products on an agarose (2%) gel. The endophyte biomass was determined as a percentage of that of the perennial ryegrass sample using the formula ($E/R \times 100$), where E and R are the quantities (pg) of endophyte and perennial ryegrass DNAs in the sample.

DNA sequencing

DNA fragments were sequenced by the dideoxynucleotide chain-termination method (Sanger et al. 1977) using Big-Dye (Version 3) chemistry with oligonucleotide primers (Sigma Genosys) specific for pUC118, pGEM-T or pGEM-T easy, and genomic sequences from *N. lolii*, *E. festucae* and *E. typhina*. Products were separated on an ABI Prism 377 sequencer (Perkin-Elmer).

Bioinformatic analyses

Sequence data were assembled into contigs using SEQUENCHER version 4.1 (Gene Codes) and analysed using the Wisconsin Package version 9.1 (Genetics Computer Group). Sequence comparisons were performed through Internet Explorer version 6.0 at the National Center for Biotechnology Information (NCBI) site (http://www.ncbi.nlm.nih.gov/) using the Brookhaven (PDB), SWISSPROT and GenBank (CDS translation), PIR and PRF databases and employing

algorithms for both local (BLASTX and BLASTP) and global (FASTA) alignments (Pearson and Lipman 1988; Altschul et al. 1990, 1997).

The nomenclature used for *E. festucae* and *L. perenne* genes follows the *A. nidulans* (Bennett and Lasure 1985) and *Oryza sativa* (Wu et al. 1991) conventions, respectively.

The following sequences have been submitted to the DDBJ/EMBL/GenBank databases: *NlggsA* (accession no. AY742904), *EfggsA* (AY742906), *EtggsA* (AY742907), *N. lolii ltm* cluster (AY742903), *E. festucaeltm* cluster (AY742905), *L. perenne* β-*Tub1* (AY742902), *L. perenne* β-*Tub2* (AY742901).

Alkaloid analyses

Lolitrem B was analysed by a modification (Panaccione et al. 2003) of the method of Gallagher et al. (1985). The amount of lolitrem B was estimated by comparison of the integrated peak areas (Class-LC10 software; Shimadzu) of the analyte with a sample of authentic lolitrem B as an external standard.

Ergovaline and peramine were analysed by previously described methods (Spiering et al. 2002; Panaccione et al. 2003). Amounts of ergovaline, together with its natural isomer ergovalinine, were estimated by comparison of the integrated peak areas (Class-LC10 software) with the sum of the peaks for ergotamine and its natural isomer ergotaminine, adjusted for relative fluorescences and extraction efficiency.

Chromatography of peramine was carried out by a column-switching procedure with an initial step to remove interfering UV absorbing compounds (Cox and Stout 1987; Spiering et al. 2002). The amount of peramine was estimated by comparison of the integrated peak area (Class-LC10 software) of the analyte with homoperamine as the internal standard.

Results

N. lolii and *E. festucae* contain two GGPP synthase genes

PCR amplification of genomic DNA from *N. lolii* (strain Lp19), *E. festucae* (strain Fl1) and *E. typhina* (strain E8) using different sets of degenerate primers based on conserved regions of domains III and V of fungal GGPP synthases resulted in single products of 209 bp (Fig. 2a) and 272 bp (Fig. 2b), respectively. The products obtained from wild-type *P. paxilli* (Fig. 2b, lane 5) correspond to sequences of the GGPP synthase genes *ggs1* (the smaller fragment) and *paxG* (the larger

fragment) (Young et al. 2001). The difference in size is due to the presence of an intron in *paxG*. Only the smaller band is present in LM662, a *paxG* deletion mutant (Fig. 2b, lane 6).

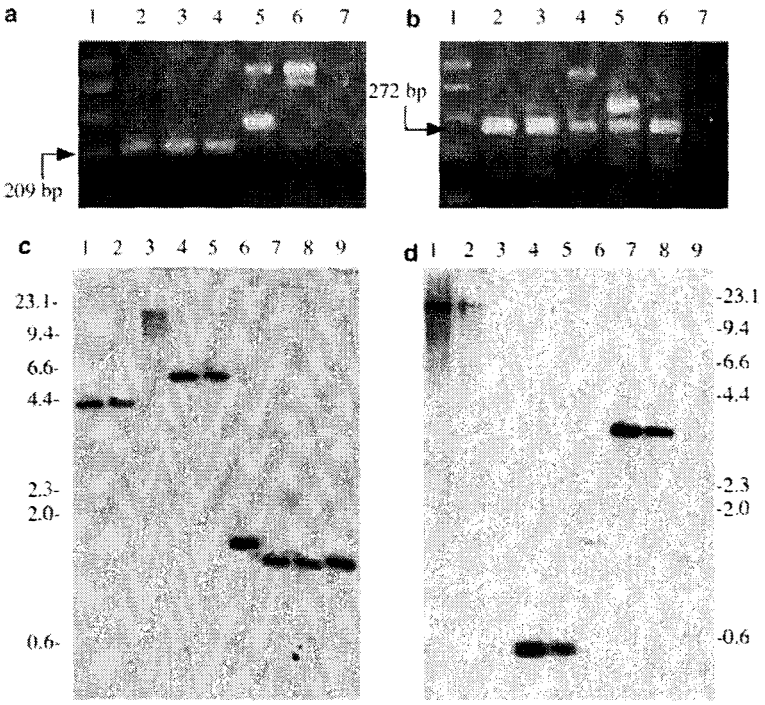


Fig. 2 a–d *N. lolii* and *E. festucae* contain two GGPP synthase genes. **a** PCR products amplified with degenerate primer set ggpps27 and ggpps28 from genomic DNAs of *N. lolii* strain Lp19 (*lane 2*), *E. festucae* strain Fl1 (*lane 3*), *E. typhina* strain E8 (*lane 4*), wild-type *P. paxilli* (*lane 5*) and *P. paxilli* strain LM662 (*lane 6*). *Lane 7* shows a water control and *lane 1* contains the 1 kb+ ladder (Invitrogen). **b** PCR products amplified with degenerate primer set ggpps27 and ggpps29 using the same genomic samples described in (**a**). **c** Autoradiograph of a Southern blot of *Eco*RI (*lanes 1–3*), *Hin*dIII (*lanes 4–6*) and *Sst*I (*lanes 7–9*) digests of genomic DNAs of *N. lolii* strain Lp19 (*lanes 1, 4 and 7*), *E. festucae* strain Fl1 (*lanes 2, 5 and 8*) and *E. typhina* strain E8 (*lanes 3, 6 and 9*) probed with the ^{32}P-labelled 272-bp insert from pCY29 (*ggs4*). λ*Hin*dIII size standards (kb) are shown on the *right*. **d** Autoradiograph of a Southern blot of the same genomic samples and digests as shown in **C** probed with the ^{32}P-labelled 209-bp insert from pCY28 (*ltmG*)

The 209- and 272-bp fragments from Lp19 were cloned into pGEM-T easy, and one clone, pCY29, containing the latter was sequenced and shown by BLASTX analysis to be most similar (*E* value of 7e−41) to the *N. crassa* GGPP synthase sequence (Accession No. AAC13867). An RFLP screen of the remaining clones identified a second unique clone, pCY28, whose sequence best matched (*E* value of 5e−19) that of the *P. paxilli ggs1* gene (Accession No. AAK11525). The two DNA sequences are 61.7% identical.

To determine which clone contained a partial sequence of the GGPP synthase specific for lolitrem biosynthesis, the plasmid inserts were amplified and used to probe a Southern blot bearing genomic digests from two lolitrem-producing strains, Lp19 and Fl1, and one non-producing strain, E8 (Fig. 2c, d). The insert from pCY29 hybridised to the DNAs of all three strains (Fig. 2c), whereas the insert from pCY28 hybridised strongly to the DNAs of the lolitrem-producing strains only (Fig. 2d). The fainter hybridisation bands observed with *E. typhina* (Fig. 2d) are identical in size to bands that hybridise in Fig. 2c, suggesting cross hybridisation to common sequences. These results indicate that the pCY29 insert is most likely to represent the gene for the "house-keeping" GGPP synthase (*ggsA*), while the insert in pCY28 is derived from the gene for the GGPP synthase (*ltmG*) that is specific for indole–diterpene (lolitrem) biosynthesis (Fig. 2).

Using the insert from pCY29 as a probe, a clone containing the full-length *ggsA*, (λCY100) was isolated from an Lp19 λGEM-12 genomic library and the complete nucleotide sequence of the gene was determined. Sequence analysis predicts that the *N. lolii ggsA* lacks introns (Fig. 3). This was confirmed by sequencing the cDNA product, amplified in a one-step RT-PCR, using as the template RNA from mycelium grown in liquid culture. To facilitate the further analysis of this gene, a comparative genomics approach was used. Orthologues of *N. lolii ggsA* were isolated from *E. festucae* Fl1 and *E. typhina* E8 using both standard and inverse PCR with primers based on the Lp19 sequence. The nucleotide sequences of *E. festucae ggsA* and *E. typhina ggsA* share 99.6 and 92% identity with *N. lolii ggsA*. Sequence analysis of *E. festucae ggsA* and *E. typhina ggsA* predicts that these genes also lack introns (Fig. 3). The 5′ translational start site of *ggsA* is difficult to define as the sequence contains three repetitive DNA elements, $CAR_{(n)}$, $GCYATCAAR_{(6)}$ and $CAG_{(n)}$ (Fig. 3). The latter is highly polymorphic within the *Epichloë*/ *Neotyphodium* group. There is no evidence for consistent intron splice sites flanking these repetitive sequences, strongly suggesting that they are coding. There are five possible in-frame start codons at the 5′ ends of the three *ggsA* genes, with the second methionine triplet having a Kozak sequence identical to the *N. crassa* consensus (CN_3CAMV**ATG**GC; Bruchez et al. 1993). A polypeptide motif (IPPRXSS) common to the N-terminal regions of the products of the *N. crassa*, *P. paxilli*, *F. fujikuroi*, *A. nidulans* and *Magnaporthe grisea ggs1* genes is also encoded by the three *ggsA* sequences, 64 bp downstream of the second methionine codon. The deduced polypeptide lengths and calculated molecular masses of NlGgsA, EfGgsA and EtGgsA, taking the second M as the start codon, are listed in Table 1.

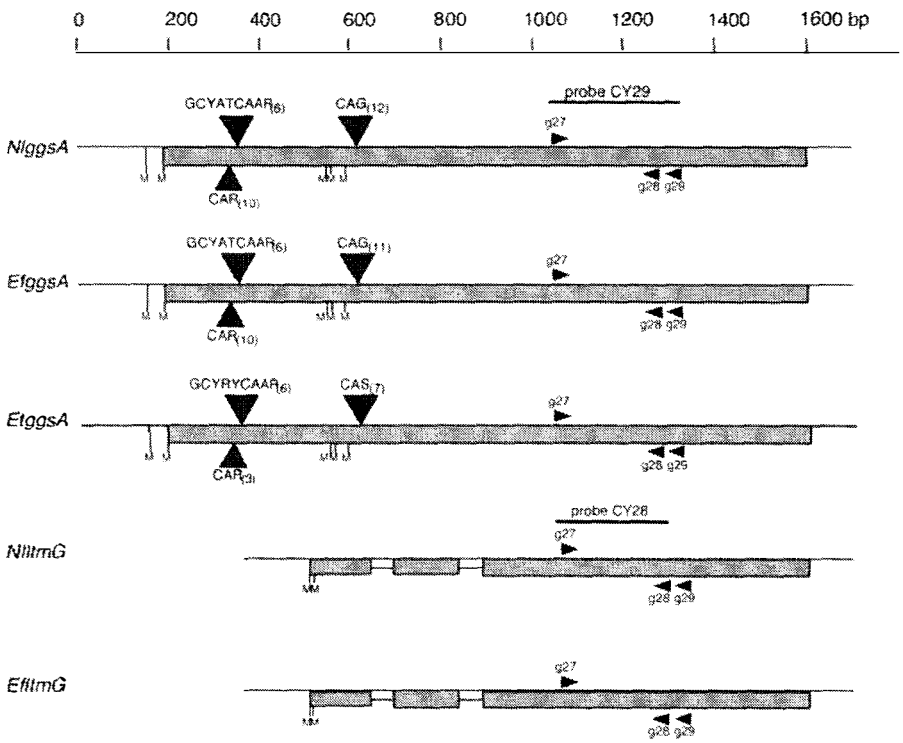


Fig. 3 Comparison of the structures of the *ggsA* and *ltmG* genes from *Epichloë* and *Neotyphodium* endophytes. The *ggsA* and *ltmG* gene structures from *N. lolii*, *E. festucae* and *E. typhina* are represented schematically. The *grey boxes* and *lines* show proposed exons and introns, respectively, in these genes. The five possible translational start sites in *ggsA* are each marked with an "M". The positions of the repeated sequences (*inverted triangles*) and primers (*arrowheads*), and the probes CY28 and CY29 (*overlines*), are also shown

---

[Table 1 will appear here. See end of document.]

To clone the full-length *N. lolii ltmG* gene the insert of pCY28 was used as a probe to isolate sequences from an Lp19 λGEM-12 genomic library. This region was found to be under-represented in the library: only five clones were isolated from ~80,000 plated. A 15.6-kb lambda clone, λCY218 (Fig. 4), was completely sequenced and shown to contain a complete copy of the *ltmG* gene. Sequence analysis of *N. lolii ltmG* predicts the presence of two introns, which were confirmed by analysis of a cDNA constructed using RNA isolated from Lp19-infected perennial ryegrass. Both introns are conserved in position relative of the three introns found in *P. paxilli paxG* (Young et al. 2001) and two of the four introns found in *F. fujikuroi ggs2* (Tudzynski and Hölter 1998). FASTA analysis showed that LtmG shares 54.1 and 52.6% identity with *N. lolii* GgsA and *P. paxilli* PaxG, respectively. The deduced polypeptide length and calculated molecular mass of LtmG is given in Table 1. LtmG contains the five conserved domains found in all prenyl diphosphate synthases (Chen et al. 1994), including the highly conserved aspartate-rich motifs (DDXXD and

DDXXN/D) in domains II and V that are proposed to serve as binding sites for the isopentenyl diphosphate (IPP) and the allyl isoprenoid substrates.

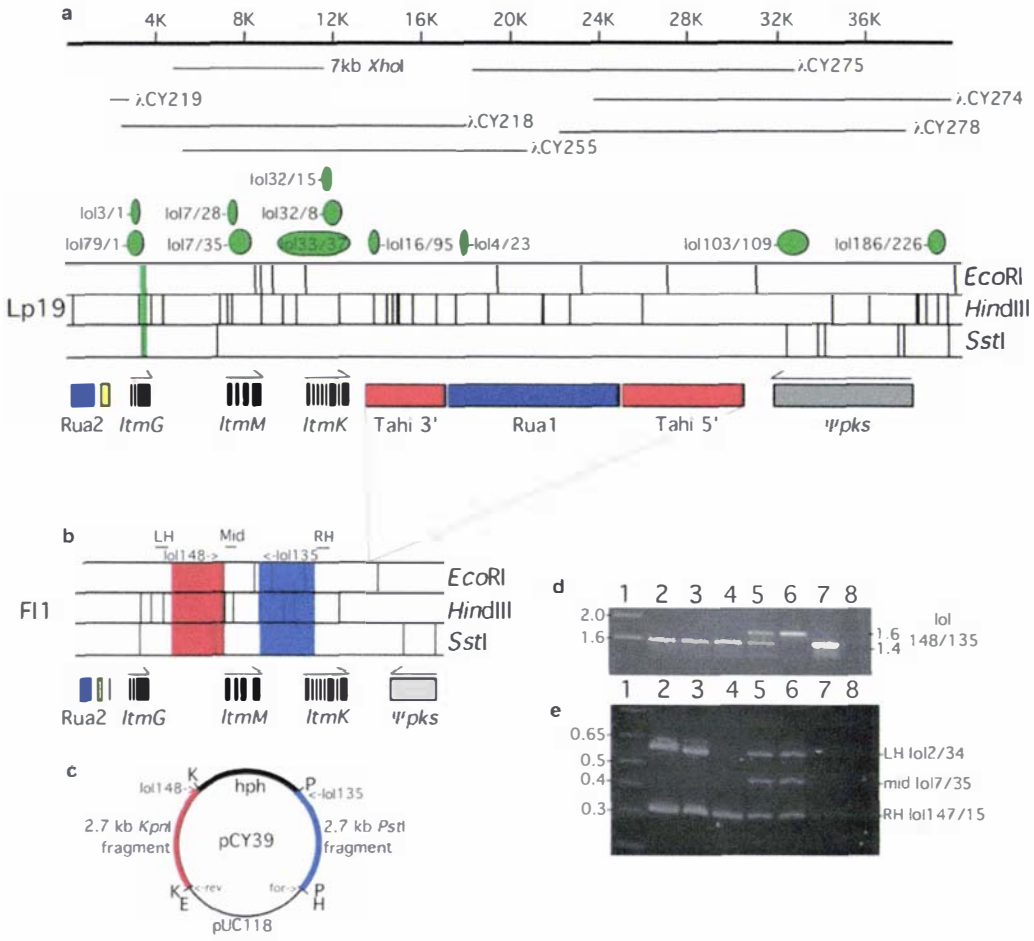


Fig. 4 a–e Physical and genetic maps of the *N. lolii* and *E. festucae ltm* loci, and construction and verification of an *ltmM* disruptant. **a, b** Physical and genetic maps of the Lp19 (**a**) and Fl1 (**b**) *ltm* loci showing the overlapping λ clones used to construct them. The position of the CY28 PCR fragment used as a probe to isolate the λ clones is highlighted by the *vertical green box*. The exons of each gene are shown as *rectangular black boxes* and the *arrows* indicate the direction of transcription. The region highlighted in *yellow* harbours a microsatellite with the core sequence TAATG. The regions highlighted in *red* and *blue* correspond to the fragments used to make the *ltmM* replacement construct. The retrotransposons, Tahi and Rua, are shown as *red* and *blue arrowed lines* with the LTR sequences indicated by the *arrowheads*. The *pks* pseudogene is *stippled* to indicate it is non-functional. Fragments used as probes in Southern hybridisations are shown as *green ovals*; the primer pairs used to amplify them are also indicated. Primer sets used to screen for targeted replacement of *ltmM* are identified by *bars* above the map in **b**. **c** Physical map of the *ltmM* replacement construct, pCY39. Primer sets used to screen for targeted replacement of *ltmM* are shown. **d** PCR screen for targeted replacement of *ltmM* in Fl1. PCR products were amplified from genomic DNA of wild-type

and mutants with the primer set indicated to the *right* of the gel. *Lane 1* 1 kb+ ladder, *lane 2* PN2303, *lane 3* PN2296, *lane 4* PN2301, *lane 5* PN2294, *lane 6* Fl1, *lane 7* pCY39, *lane 8* H_2O control. **e** PCR screen for targeted replacement of *ltmM* in Fl1. Samples are identical to those shown in (**d**) but PCRs was performed with the primer sets indicated

---

N. lolii and *E. festucae* both contain a cluster of genes for indole–diterpene biosynthesis

Bioinformatic analysis of the remaining ltm genes are flanked by AT rich DNA

Bioinformatic analysis of the remaining λCY218 sequence identified two additional genes with significant sequence similarity to genes from *P. paxilli* known to be required for paxilline biosynthesis (Table 1). *N. lolii ltmM* and *N. lolii ltmK* are homologues of *paxM* (E value 5e-96), a FAD-dependent monooxygenase, and *paxP* (E value of 6e-63), a cytochrome P450 monooxygenase, respectively. The corresponding genome region was sequenced from *E. festucae* strain Fl1 and shown to share 99.9% identity, from the start of *ltmG* to the stop codon of *ltmK*, with the Lp19 nucleotide sequence (Fig. 4). The polypeptide sequence of *N. lolii lltmG* is identical to that of *E. festucae ltmG*.

Sequence analysis of cDNA products amplified from RNA isolated from Lp19-infected perennial ryegrass confirmed that *ltmM* has three introns. The positions and phase of the first two introns are conserved with those found in *P. paxilli paxM*. A comparison of the Lp19 and Fl1 sequences identified just two nucleotide differences. Both were transitions: an A to G at position 91 and a T to C at position 249 (position 1 being the A of the ATG start codon). Only the first results in an amino acid residue change with a conservative replacement of methionine (in Lp19) to valine (in Fl1). The promoter region of *E. festucae ltmM* has two differences to that of *NlltmM*: the first, a T to C transition at position -356 results in the loss of a *Hin*dIII site, and the second an insertion of GA at position −1038. The deduced polypeptide length and unmodified molecular mass of *LtmM* is shown in Table 1. The overall sequence identity of *N. lolii LtmM* to *PaxM* is 41.0%. ClustalW alignment of *LtmM* with *PaxM* and other closely related polypeptide sequences, identifies the presence of four highly conserved motifs: the dinucleotide binding domain (Wierenga et al. 1996), the ATG motif (Vallon 2000), a GD motif (Eggink et al. 1990) and a G-helix. These motifs are good indicators of a modified Rossman fold, used by many flavoproteins to bind FAD.

Sequence analysis of LtmK indicates it is more similar to PaxP than PaxQ sharing 31.3% and 23.4% identity, respectively. Despite the relatively high identity (E value of 6e-63) between LtmK and PaxP at least three cytochrome P450 enzymes are predicted to be required for lolitrem B biosynthesis (see Discussion). Indeed, two additional cytochrome P450 genes with greater homology to *paxP* and *paxQ* have been identified in a suppression substractive hybridization screen (R Johnson, personal communication). For this reason the gene has been named *ltmK* rather than *ltmP*. Characterisation of the *ltmK* sequence by cDNA analysis identified seven introns, four (introns 1, 2, 4 and 7) of which are conserved in position with *paxP* and three (introns 2, 3 and 8) conserved in position with *paxQ*. *N. lolii* LtmK contains the classical signature motifs of cytochrome P450 enzymes, including a haem-binding domain (**FG**Y**G**TWA**C**P**G**RF**LA**) (Graham-Lorence 1996). The deduced polypeptide length and unmodified molecular mass of LtmK is shown in Table 1. The amino acid sequence of *E. festucae* LtmK is identical to that of *N. lolii* LtmK.

ltm genes are flanked by AT rich DNA that has undergone extensive repeat-induced point mutation (RIP)

To obtain further sequence to the left of *ltmG*, the Lp19 λGEM-12 library was screened with a probe amplified with primer set lol3 and lol1. Analysis of the hybridising clones identified one clone, λCY219, which contained additional flanking sequence (Fig. 4). However, this clone was severely rearranged and only 1,051 bp corresponded to the correct genomic sequence. Given the instability of library clones in this region, further sequence (1.9 kb) on this side was obtained by inverse PCR. Analysis of this Lp19 sequence identified a large microsatellite, with a $(TAATG)_{101}$ consensus sequence. The same microsatellite is present in Fl1, but here it is interrupted by a 161-bp insertion, resulting in separate blocks of $(TAATG)_{67}$ and $(TAATG)_{22}$ (Fig. 4). The genome sequence to the left of the microsatellite has an AT content of 82.5%.

The DNA sequence flanking the right-hand end of the *ltm* gene cluster is also very AT-rich (71.2%) compared to the composition of the *ltm* (59.3%) and *ggsA* (40.9%) genes. BLASTX analysis of this region indicates that this sequence is of retrotransposon origin but is highly corrupted, with no evidence of intact ORFs. Southern analysis of genomic digests of Lp19, Fl1 and E8 with a probe generated with primer set lol4 and lol23 (Fig. 4) revealed that these AT-rich sequences are an abundant and highly dispersed component of the *N. lolii* and *E. festucae* genomes, but are absent from the genome of *E. typhina* (Fig. 5b).

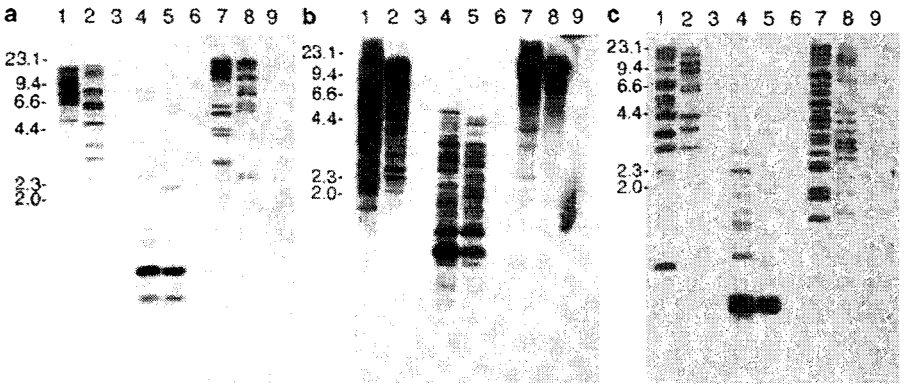


Fig. 5 a–c Retrotransposon sequences are abundant and dispersed components of the *N. lolii* and *E. festucae* genomes. **a** Southern analysis of *N. lolii* strain Lp19 (*lanes 1, 4 and 7*), *E. festucae* strain Fl1 (*lanes 2, 5 and 8*) and *E. typhina* strain E8 (*lanes 3, 6 and 9*). Genomic digests were probed with a ^{32}P-labelled Tahi fragment generated with primer set lol16 and lol95. *Lanes 1–3 Eco*RI digest, *lanes 4–6 Hin*dIII and *lanes 7–9 Sst*I. The sizes of standards are in kb. **b** Autoradiograph of the same blot as shown in (**a**), probed with a ^{32}P-labelled Rua fragment generated using primer set lol4 and lol23. **c** Autoradiograph of the same blot as in (**a**), probed with a ^{32}P-labelled 787-bp fragment generated using primer set lol186 and lol226

To obtain additional sequence to the right of *ltmK*, two approaches were used. The first involved screening the Lp19 λGEM-12 library with an *ltmK* probe that was generated with primers lol33 and lol37 (Fig. 4). This led to the isolation of a clone, λCY255, which overlaps with λCY218 (Fig. 4). The second approach was to use inverse PCR to "hop" to the other side of the "retro-platform", using Fl1, a strain shown by Southern analysis to lack this particular retro-platform. Using a probe generated with primer set lol103 and lol109, the Lp19 clones λCY275, λCY278 and λCY274 were isolated from a λGEM-12 library. Clone λCY275 was found to overlap with λCY255 (Fig. 4). Southern analysis, using genomic DNAs digested with *Sst*I and *Sal*I, recognition sites for which are absent in the Lp19 retro-platform, confirmed that the only difference in genome organisation between Lp19 and Fl1 is the presence of the retro-platform in Lp19 (data not shown).

Bioinformatic analysis of the Lp19 retro-platform revealed it is comprised of two independent retro-elements: one, which we have named Tahi (Māori for "one") contains an insertion of the other, which we have named Rua (Māori for "two") (Fig. 4). Domains of the retrotransposon *pol* gene, including protease, reverse transcriptase, RNaseH and integrase, were recognisable in both elements. However, the AT bias of this sequence introduces many stop codons. No recognisable *gag* motifs were found in either element. The identification of long terminal repeats (LTRs) of approximately 104 and 272 bp, and target-site duplications of 5 bp flanking both elements, suggests that Rua inserted into Tahi. A second partial copy of Rua (Rua-2) is located upstream of the

microsatellite adjacent to *ltmG* (Fig. 4). Rua-2 contains one identifiable LTR and a partial *pol* sequence, but lacks any *gag* sequence.

The genomic abundance of the Tahi element was determined by Southern analysis (Fig. 5a). This retrotransposon is present in fewer copies in the genomes of *N. lolii* and *E. festucae* than is the Rua element (see above) and is absent from the *E. typhina* genome.

The AT content of the Tahi and Rua elements is 72.2 and 70.0%, respectively. Sequence alignment of the three Rua and two Tahi LTR sequences, together with other retrotransposon sequences obtained from a library screen with a probe generated from primers lol4 and lol23, revealed a strong bias for GC-to-AT mutations. This mutational bias is usually associated with repetitive sequences that have undergone a RIP event (Selker et al. 1987; Cambareri et al. 1989).

Bioinformatic analysis of λCY278 and λCY274 sequence data identified a polyketide synthase (PKS) pseudogene adjacent to the retro-platform, flanked on the right by an additional AT-rich (78%) sequence. PCR analysis of the corresponding Fl1 sequence confirmed the presence of a PKS pseudogene, containing many frame-shift mutations, immediately adjacent to *ltmK*. Southern analysis using a probe for the AT-rich region flanking the PKS pseudogene confirmed that this sequence is also present in multiple copies in the genomes of Lp19 and Fl1 (Fig. 5c).

ltmM is essential for lolitrem B biosynthesis

To confirm that *ltmM* is essential for lolitrem biosynthesis, a replacement construct, pCY39, was prepared and a PCR-generated linear component of this plasmid was recombined into the genome of *E. festucae* strain Fl1. Potential *ltmM* replacements were identified by PCR screening of 159 nuclear-purified Hyg^R transformants using a primer pair (lol148 and lol135) that flanks the *ltmM* region to be deleted (Fig. 4). In this first screen 9/159 (5.7%) of the transformants yielded a single 1.4-kb PCR product (Fig. 4d, lanes 2–4), which is consistent with a targeted replacement event. By contrast, ectopic transformants give two PCR products (Fig. 4d, lane 5), a 1.6-kb wild-type fragment (Fig. 4d, lane 6) and a 1.4-kb *hph* fragment (Fig. 4d, lane 7). The nine replacement mutants were subjected to a second PCR screen using primer sets that amplify upstream (lol2 and lol34: 574 bp), internal (lol7 and lol35: 448 bp) and downstream (lol147 and lol15: 317 bp) regions of *ltmM* (Fig. 4e). All nine mutants were confirmed to be *ltmM* replacements (Fig. 4e, lanes 2 and 3) but in strain PN2301 the deletion extended to the left to an undefined point beyond *ltmG* and the flanking retro-platform (Fig. 4e, lane 4). Southern analysis of *Bam*HI and *Eco*RI digests of these mutant DNAs probed with *ltmM* (lol7/lol35) and *hph* (pUChph3/ pUChph4) probes, respectively, revealed that five (3.1% of total) of the nine mutants were true replacements. The

other four mutants had, in addition to the internal deletion, single or multiple copies of the construct integrated either in the junction sequences or in ectopic sites (data not shown).

To determine the chemical and symbiotic phenotype of these mutants, perennial ryegrass seedlings were infected with two *ltmM* deletion mutants (PN2303 and PN2296) an extended *ltmMG* deletion mutant (PN2301), an ectopic transformant (PN2294), and wild-type F11. Two months after infection, seedlings were examined for the presence of the endophyte by immunoblotting. There were no differences in infection rates or in hyphal morphology and growth between mutant and wild-type strains (results not shown). Using real-time PCR to quantitate fungal *ggsA* (endo1/endo2) and plant β-*Tub1* (PRG12/PRG13) DNA, the relative endophyte biomass, as estimated from the relative DNA content of each partner, was within the range of 0.3–1.9% of the total biomass for both wild-type and mutant associations (Table 2). Examination of alkaloid levels in the plant material showed that plants containing endophyte strains deleted for *ltmM* (PN2303, PN2296, PN2301) were unable to produce lolitrem B (Fig. 6). Two minor peaks, with elution profiles consistent with lolitrems A and E (Miles et al. 1994; Munday-Finch et al. 1995), were also absent in extracts of plants infected with these three mutants. By contrast, extracts of plant material infected with PN2294, an ectopic transformant, had levels of all three lolitrems comparable to wild-type (Fig. 6). The levels of ergovaline and peramine, the other major alkaloids known to be produced by these endophyte strains, were not significantly different between wild-type and mutants (Table 2).

[Table 2 will appear here. See end of document.]

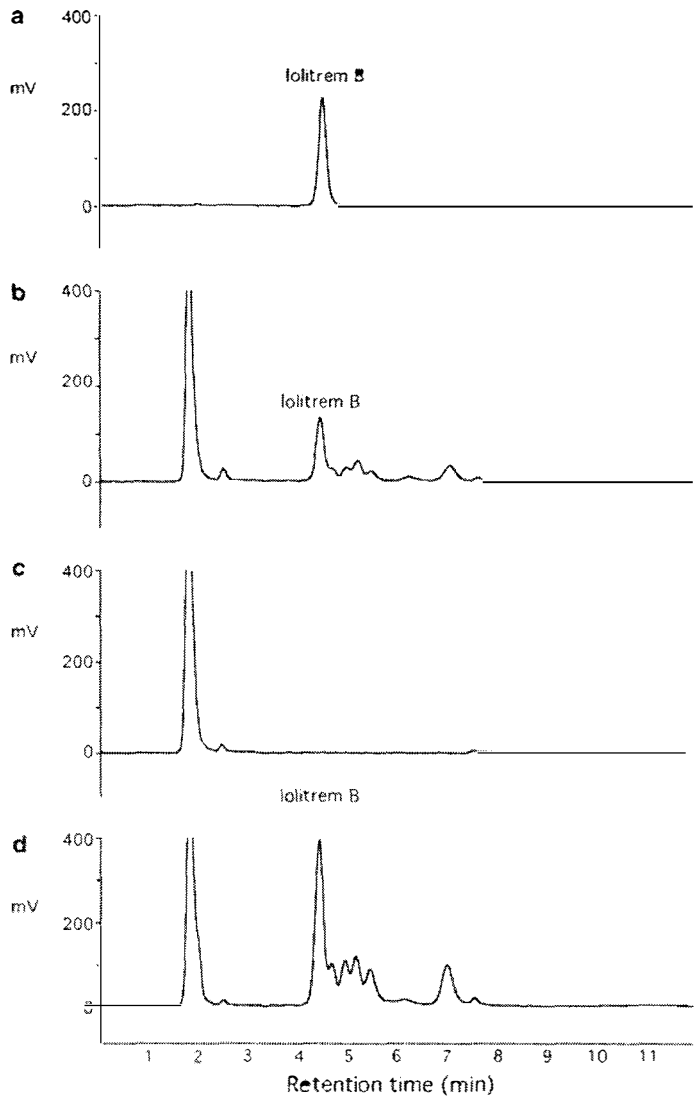


Fig. 6 **a–d** HPLC analysis of lolitrem alkaloids in leaf extracts of endophyte infected perennial ryegrass. Pseudostem tissue was harvested 2 months post-infection and analysed for lolitrems by normal phase HPLC. **a** Lolitrem B standard (8.4 μg). **b** Wild-type strain F11 (plant G1137). **c** *ltmM* mutant PN2303 (plant G1114). **d** Ectopic transformant PN2294 (plant G1130). The *y*-axis shows fluorescence units in millivolts (emission wavelength 440 nm) and the *x*-axis indicates retention time in min. The peak at 1.9 min corresponds to the solvent front

To confirm that deletion of *ltmM* was responsible for the lolitrem-negative phenotype, the construct pCY40, containing a wild-type copy of *ltmM*, was introduced into the mutants PN2303 and PN2301, deleted for *ltmM*. An arbitrary selection of these transformants, together with controls containing the vector only, was used to infect perennial ryegrass seedlings, and the lolitrem phenotype of these associations determined (Table 3). This analysis confirmed that pCY40 was

able to complement PN2303 (Table 3) but was unable to complement PN2301 (results not shown). The latter result confirms that at least one other gene besides *ltmM* is required for lolitrem biosynthesis. A complementation test was also carried out using pCY41, a construct containing *paxM*, the *P. paxilli* orthologue of *ltmM*. Of three arbitrarily selected transformants tested, one strain, PN2352, was complemented by pCY41. Southern analysis of this and the two other transformants tested showed that PN2352 has a large number of copies of pCY41 integrated into the genome of PN2303, whereas the strains that were not complemented had only a few copies of pCY41 (results not shown). This high copy number is proposed to be the reason why this transformant was able to synthesise lolitrem B in planta.

[Table 3 will appear here. See end of document.]

ltm genes are preferentially and highly expressed in planta

Northern analysis of *N. lolii* and *E. festucae ltm* gene expression demonstrated that *ltmG*, *ltmM* and *ltmK* are very weakly expressed in culture but strongly induced in planta (Fig. 7). Taking into account the fact that endophyte biomass represents about 1% of the plant biomass (see previous section) the steady-state levels of *ltmG*, *ltmM* and *ltmK* transcripts are about 100-fold greater than the levels of the perennial ryegrass *Act1* transcript. In contrast, the endophyte *tubB* transcript is not detectable in either Lp19- or F11-infected perennial ryegrass, but is detectable in mycelium grown in culture. To improve the sensitivity of the analysis, the expression of *N. lolii ltm* genes was examined using reverse transcriptase-PCR (Fig. 8). In RNA isolated from mycelium grown under a range of physiological conditions, transcripts of *ltmG* and *ltmM* could be detected in amplifications starting from a 1/10 dilution of the cDNA (Fig. 8a). By contrast, *ltmK* was detectable in amplifications of a 1/100 dilution of the cDNA, which was comparable to the steady-state level of *tubB*. Using RNA isolated from Lp19-infected perennial ryegrass, levels of *ltmG*, *ltmM* and *ltmK* transcripts were similar to those of the perennial ryegrass β-*Tub1* RNA (Fig. 8b), a result which is consistent with the Northern analysis. This method showed that transcripts of the endophyte *tubB* gene were detectable in plant material, but only in amplifications of a 1/10 dilution of the cDNA. Taken together, these results indicate that the *ltmG*, *ltmM* and *ltmK* genes of *N. lolii* and *E. festucae* are very highly expressed in planta, but only weakly expressed in culture. To determine whether the very weak expression of *ltm* genes in culture was due to either carbon or nitrogen repression, reverse transcriptase-PCR analysis was carried out using RNA isolated from cultures grown under conditions of carbon and nitrogen limitation (Fig. 8a: columns 2–4). While a trace of transcript could be detected in amplifications of some of these cDNA samples, there were no

noticeable differences in levels of transcripts between these cultures and cultures grown under conditions of carbon and nitrogen excess (Fig. 8a: columns 5–7).

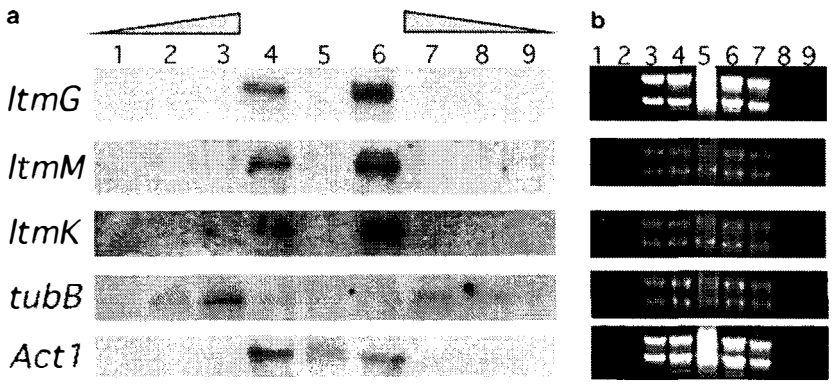


Fig. 7 a, b Northern analysis of *N. lolii* and *E. festucae ltm* gene expression. **a** Northern analysis of RNAs isolated from mycelia of Lp19 (*lanes 1–3*) or Fl1 (*lanes 7–9*), and pseudostems of perennial ryegrass (Nui) infected with Lp19, not infected or infected with Fl1 (*lanes 4–6*), probed with ^{32}P-labelled Fl1 *ltmG*, *ltmM*, *ltmK*, *tubB* cDNAs and perennial ryegrass *Act1* cDNA. **b** The ethidium bromide stained 16S and 26S rRNA region of each gel is shown for comparison. *Lane 1* 0.15 µg Lp19 RNA, *lane 2* 1.5 µg Lp19 RNA, *lane 3* 15 µg Lp19 RNA, *lane 4* 15 µg of Lp19 infected Nui RNA, *lane 5* 15 µg uninfected Nui RNA, *lane 6* 15 µg of Fl1 infected Nui RNA, *lane 7* 15 µg Fl1 RNA, *lane 8* 1.5 µg Fl1 RNA and *lane 9* 0.15 µg Fl1 RNA

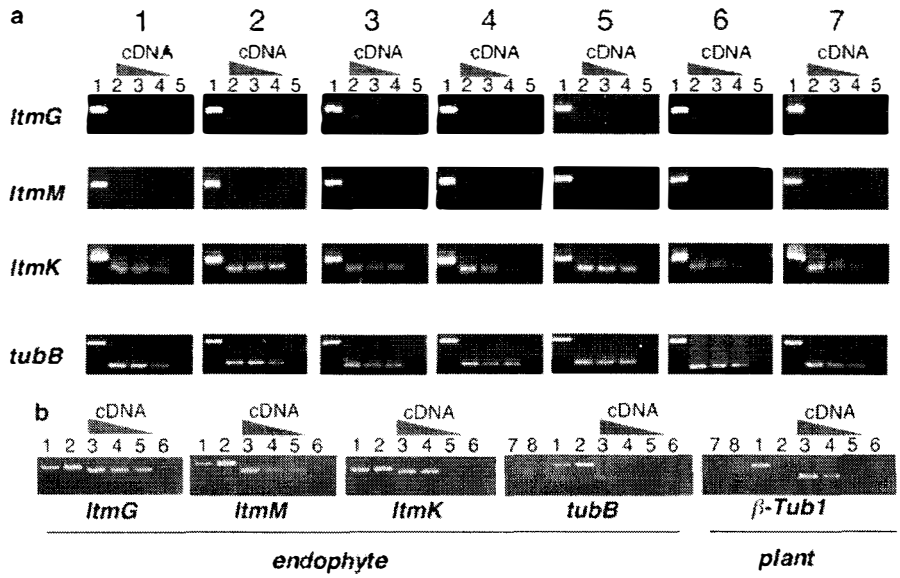


Fig. 8 a, b RT-PCR analysis of *N. lolii ltm* gene expression. **a** Reverse transcriptase-PCR analysis of steady-state levels of *ltm* and *tubB* transcripts in Lp19 mycelium. As templates 1/10 (*lanes 2*), 1/100 (*lanes 3*) and 1/1,000 (*lanes 4*) dilutions of the cDNA were used. *Lane 5* shows the water only control, *lane 1* the product obtained from genomic DNA. RNA was isolated from mycelia grown in liquid culture in PD (1).

CD salts (2), CD salts plus 1% glycerol (3), CD salts plus 1% glycerol and 100 mM NH_4Cl (4), CD salts plus 100 mM NH_4Cl (5), CD salts plus 10% glucose (6), CD salts plus 10% glucose and 100 mM NH_4Cl (7). **b** Reverse transcriptase-PCR analysis of Lp19 *ltm* and *tubB* transcripts, and the perennial ryegrass β-*Tub1* transcript, in Lp19-infected perennial ryegrass pseudostem tissue. As templates 1/10 (*lanes 3*), 1/100 (*lanes 4*), 1/1,000 (*lanes 5*) dilutions were used. *Lane 6* shows the water only control. The products amplified from DNA obtained from perennial ryegrass infected with Lp19 (*lane 1*) or from Lp19 mycelium (*lane 2*) are also shown. *Lanes 7 and 8* are controls containing no reverse transcriptase and no RNA, respectively

Discussion

We describe here the molecular cloning and genetic analysis of a cluster of three symbiotically regulated genes from *N. lolii* and *E. festucae* that are required for synthesis of the plant bioprotective metabolite lolitrem B. All three genes are weakly expressed in culture but strongly expressed in planta. Using quantitative real-time PCR, the endophyte biomass in these associations was estimated to be between 0.3 and 1.9% of the total biomass. Taking into account this biomass difference, the steady-state levels of *ltmG*, *ltmM* and *ltmK* are about 100-fold greater than the levels of the perennial ryegrass β-*Tub1* and *Act1* transcripts. The fact that these genes are not de-repressed in culture in response to low levels of either nitrogen or carbon suggests that their regulation is plant-specific rather than the result of a response to low carbon or nitrogen levels. There are now several reports of plant-induced fungal genes, including the *mig* genes of *Ustilago maydis* (Basse et al. 2002) and the *ACE1* and *PLS1* genes of *Magnaporthe grisea* (Clergeot et al. 2001; Bohnert et al. 2004), but, to date, the signalling mechanism has not been determined.

Endophyte synthesis of lolitrems is proposed to confer major fitness benefits on the grass host by protecting it from herbivory by insects and other small animals (Clay 1990; Clay and Schardl 2002). The *ltm* cluster in Lp19 is atypical of most fungal gene clusters for biosynthesis of secondary metabolites reported to date (Brown et al. 2001; Young et al. 2001; Bradshaw et al. 2002; Bhatnagar et al. 2003; Proctor et al. 2003) in that it is embedded in a retrotransposon-rich region of the genome, although a recent report (Gardiner et al. 2004) suggests that the sirodesmin biosynthetic gene cluster in the plant pathogenic fungus *Leptosphaeria maculans* is also associated with a retrotransposon-rich region of the genome. The right-hand side of the *ltm* cluster in Lp19 is flanked by a 17.2-kb retrotransposon "platform" comprised of one retro-element (Rua) inserted within another (Tahi). The presence of LTRs flanking both elements, 5-bp direct repeats at the target sites, and highly corrupted remnants of a *pol* gene encoding a protease, reverse transcriptase,

RNaseH, and integrase, in that order, would suggest that both elements are members of the Class I gypsy sub-family of retrotransposons (Kempken and Kück 1998). Both elements are abundant and highly dispersed components of the genomes of *E. festucae* and *N. lolii*. Neither of these elements was present in the genome of *E. typhina*, a species that is able to form asexual inter-specific hybrids with *E. festucae* (Schardl et al. 1994). Adjacent to the right-hand retro-platform is an inactive polyketide synthase gene, which is flanked by another region of AT-rich DNA. On the left-hand side of the *ltm* cluster is a retrotransposon comprised of at least one Rua element. The instability of genomic clones isolated from this region and the highly repetitive nature of the DNA sequence has made it extremely difficult to rescue additional sequence on this side of the *ltm* cluster. The high frequency of GC-to-AT mutations within these elements suggests that these sequences were previously subjected to RIP, a pre-meiotic gene silencing process first identified in *N. crassa* (Selker et al. 1987; Cambareri et al. 1989). Whether RIP is still active in *E. festucae*, as has recently been found in some other sexual species of filamentous fungi (Idnurm and Howlett 2003), remains to be determined. The abundance of both the Tahi and Rua classes of retro-elements in both *N. lolii* and *E. festucae* indicates that acquisition of RIP preceded the divergence of the asexual from the sexual species. The absence of these elements in *E. typhina* could indicate that these elements were acquired after the divergence of the *Epichloë* species.

The lolitrems are closely related to paxilline (Fig. 1), an abundant metabolite synthesised by cultures of *P. paxilli*. The isolation from *P. paxilli* of a cluster of genes required for paxilline biosynthesis has provided important insights into the biochemistry of indole-diterpene biosynthesis (Young et al. 2001; McMillan et al. 2003; Parker and Scott 2004). A comparison with the paxilline biosynthesis cluster suggests that *ltmG* and *ltmM* are functional orthologues of *paxG* and *paxM*. The third gene in the cluster, *ltmK*, does not appear to have an orthologue in the *pax* gene cluster, and is therefore proposed to be unique to lolitrem biosynthesis.

The first of these genes, *ltmG*, is predicted to encode a GGPP synthase, which is proposed to catalyse the determinant step for lolitrem biosynthesis, i.e., the synthesis of GGPP. Interestingly, the three fungal species in which diterpene gene clusters have been extensively analysed each have two GGPP synthases, one for primary metabolism and one that is specifically recruited for secondary metabolism (Tudzynski and Hölter 1998; Young et al. 2001; Zhang et al. 2004). The *N. lolii* and *E. festucae* strains analysed here also have two GGPP synthases, whereas the lolitrem negative strain of *E. typhina* that was analysed here, strain E8, had only a single gene for GGPP synthase. The 5′ region of the GGPP synthase gene for primary metabolism, *ggsA*, was found to be highly polymorphic in the three endophyte species analysed, lacked introns and was weakly

expressed both in culture and in planta. The GGPP synthase gene for secondary metabolism, *ltmG*, found in the lolitrem producing strains of *N. lolii* and *E. festucae*, appears to be an orthologue of *paxG* from *P. paxilli ggs-2* from *F. fujikuroi*, and *atmG* from *A. flavus*—genes that are required for the biosynthesis of the diterpenes paxilline, gibberellins and aflatrems, respectively (Tudzynski and Hölter 1998; Young et al. 2001; Zhang et al. 2004). Deletion and complementation analysis of *ltmM* established that this gene is essential for lolitrem biosynthesis. Complementation of the *ltmM* mutant with a copy of the *P. paxilli paxM* confirms that *ltmM* is a functional orthologue of *paxM*. By contrast, *ltmM* under the control of its native promoter is unable to complement a *paxM* deletion mutant of *P. paxilli* (C. A. Young and B. Scott, unpublished results). While deletion of *paxM* in *P. paxilli* results in a paxilline-negative phenotype, to date, no identifiable indole–diterpene intermediate has been identified in this mutant (B. Scott, L. K. McMillan, J. W. Astin, S. Saikia, C. A. Young, A. Bryant and E. J. Parker, unpublished results). This would suggest that *paxM* is involved in a very early step in the pathway. Our working model is that PaxM is required to catalyse the epoxidation of GGPP, which is subsequently cyclised by PaxC to form the first stable indole–diterpene, possibly paspaline, after the addition of indole-3-glycerol phosphate (Parker and Scott 2004). By analogy we propose *ltmM* catalyses the same early step in lolitrem biosynthesis. An *N. lolii* or *E. festucae* orthologue of *paxC* has yet to be identified but, based on genetic analysis of *paxC* in *P. paxilli*, is predicted to be essential for lolitrem biosynthesis (Parker and Scott 2004).

Other genes essential for paxilline biosynthesis are *paxP* and *paxQ*; which encode cytochrome P450 enzymes (McMillan et al. 2003). PaxP is proposed to be required for demethylation of C-12 of paspaline and possibly for hydroxylation of C-10, and PaxQ is required for hydroxylation of C-13 using either PC-M6 or 13-desoxypaxilline as substrates. Analysis of the structure of lolitrem B (Fig. 1) suggests that similar modifications of paspaline are required to form lolitrem B. *N. loli* or *E. festucae* orthologues of *paxP* and *paxQ* have yet to be identified. One or two additional cytochrome P450 enzymes are required to complete the oxidation and closure of ring-A of lolitrem B. LtmK is a recognisable cytochrome P450 monooxygenase and may catalyse these or other oxidation steps unique to lolitrem biosynthesis. Additional oxidation and prenylation steps are required to form ring I of lolitrem B. In summary, we predict that at least ten genes are required for the biosynthesis of lolitrem B, the three described here, plus additional genes to carry out the functions described above.

The cloning of the *ltmG– tmM– ltmK* cluster of genes will now allow us to study how these genes are regulated in the plant and whether their expression alters in response to the imposition of various biotic or abiotic stresses on the symbiotum. We will also be able to investigate how

these and the other *ltm* genes are organised in the genome, and how this cluster has evolved. Importantly, by a combination of comparative genomics and genetic analysis in both *E. festucae* and *P. paxilli*, we will be able to elucidate the biochemistry of lolitrem biosynthesis.

Acknowledgements This research was supported by grants MAU-X0127 and C10X0203 from the New Zealand Foundation for Research, Science and Technology (FRST), and a grant (MAU103) from the Royal Society of New Zealand Marsden Fund. The authors thank Andrea Bryant (Massey) for technical assistance, Joanne Dobson for constructing the Lp19 genomic library, Wayne Simpson and Elizabeth Davies (AgResearch) for technical assistance and advice, and Emily Parker (Massey) for discussions on the chemistry of lolitrem B biosynthesis. Carolyn Young was a recipient of a FRST Bright Futures Scholarship.

References

Acklin W, Weibel F, Arigoni D (1977) Zur Biosynthese von Paspalin und verwandten Metaboliten aus *Claviceps paspali*. Chimia 31:63

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Basse CW, Kolb S, Kahmann R (2002) A maize-specifically expressed gene cluster in *Ustilago maydis*. Mol Microbiol 43:75–93

Bennett JW, Lasure LL (1985) Conventions for gene symbols. In: Bennett JW, Lasure LL (eds) Gene manipulation in fungi. Academic, London, pp 537–544

Bhatnagar D, Ehrlich KC, Cleveland TE (2003) Molecular genetic analysis and regulation of aflatoxin biosynthesis. Appl Microbiol Biotechnol 61:83–93

Bohnert HU, Fudal I, Dioh W, Tharreau D, Notteghem JL, Lebrun MH (2004) A putative polyketide synthase/peptide synthetase from *Magnaporthe grisea* signals pathogen attack to resistant rice. Plant Cell 16:2499–2513

Bradshaw RE, Bhatnagar D, Ganley RJ, Gillman CJ, Monahan BJ, Seconi JM (2002) *Dothistroma pini*, a forest pathogen, contains homologs of aflatoxin biosynthetic pathway genes. Appl Environ Microbiol 68:2885–2892

Brown DW, McCormick SP, Alexander NJ, Proctor RH, Desjardins AE (2001) A genetic and biochemical approach to study trichothecene diversity in *Fusarium sporotrichioides* and *Fusarium graminearum*. Fungal Genet Biol 32:121–133

Bruchez JJP, Eberle J, Russo VEA (1993) Regulatory sequences involved in the translation of *Neurospora crassa* mRNA: Kozak sequences and stop codons. Fungal Genet Newslett 40:85–88

Bullock WO, Fernandez JM, Short JM (1987) XL1-Blue: a high efficiency plasmid transforming *recA*Escherichia coli strain with beta-galactosidase selection. Biotechniques 5:376–378

Bush LP, Wilkinson HH, Schardl CL (1997) Bioprotective alkaloids of grass-fungal endophyte symbioses. Plant Physiol 114:1–7

Byrd AD, Schardl CL, Songlin PJ, Mogen KL, Siegel MR (1990) The β-tubulin gene of *Epichloëtyphina* from perennial ryegrass (*Lolium perenne*). Curr Genet 18:347–354

Byrne KM, Smith SK, Ondeyka JG (2002) Biosynthesis of nodulisporic acid A: precursor studies. J Am Chem Soc 124:7055–7060

Cambareri EB, Jensen BC, Schabtach E, Selker EU (1989) Repeat-induced G-C to A-T mutations in Neurospora. Science 244:1571–1575

Carroll AM, Sweigard JA, Valent B (1994) Improved vectors for selecting resistance to hygromycin. Fungal Genet Newslett 22:

Chen AP, Kroon PA, Poulter CD (1994) Isoprenyl diphosphate synthases: protein sequence comparisons, a phylogenetic tree, and predictions of secondary structure. Protein Sci 3:600–607

Christensen MJ, Leuchtmann A, Rowan DD, Tapper BA (1993) Taxonomy of *Acremonium* endophytes of tall fescue (*Festuca arundinacea*), meadow fescue (*F. pratensis*) and perennial rye-grass (*Lolium perenne*). Mycol Res 97:1083–1092

Clay K (1990) Fungal endophytes of grasses. Annu Rev Ecol Syst 21:275–297

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

Clergeot PH, Gourgues M, Cots J, Laurans F, Latorse MP, Pepin R, Tharreau D, Notteghem JL, Lebrun MH (2001) *PLS1*, a gene encoding a tetraspanin-like protein, is required for penetration of rice leaf by the fungal pathogen *Magnaporthe grisea*. Proc Natl Acad Sci USA 98:6963–6968

Cole RJ, Dorner JW, Lansden JA, Cox RH, Pape C, Cunfer BM, Nicholson SS, Bendell DM (1977) Paspalum staggers: isolation and identification of tremorgenic metabolites from sclerotia of *Claviceps paspali*. J Agric Food Chem 25:1197–1201

Cox GB, Stout RW (1987) Study of the retention mechanisms for basic compounds on silica under "pseudo-reversed phase" conditions. J Chromatogr 384:315–336

De Jesus AE, Gorst-Allman CP, Steyn PS, van Heerden FR, Vleggar R, Wessels PL, Hull WE (1983) Tremorgenic mycotoxins from *Penicillium crustosum*. Biosynthesis of Penitrem A. J Chem Soc Perkin Trans 1863–1868

Fletcher LR, Harvey IC (1981) An association of a *Lolium* endophyte with ryegrass staggers. NZ Vet J 29:185–186

Frischauf AM, Lehrach H, Poustka A, Murray N (1983) Lambda replacement vectors carrying polylinker sequences. J Mol Biol 170:827–842

Gallagher RT, White EP, Mortimer PH (1981) Ryegrass staggers: isolation of potent neurotoxins lolitrem A and lolitrem B from staggers-producing pastures. NZ Vet J 29:189–190

Gallagher RT, Campbell AG, Hawkes AD, Holland PT, McGaveston DA, Pansier EA (1982) Ryegrass staggers: the presence of lolitrem neurotoxins in perennial ryegrass seed. NZ Vet J 30:183–184

Gallagher RT, Hawkes AD, Steyn PS, Vleggaar R (1984) Tremorgenic neurotoxins from perennial ryegrass causing ryegrass staggers disorder of livestock: structure elucidation of lolitrem B. J Chem Soc Chem Commun 614–616

Gallagher RT, Hawkes AD, Stewart JM (1985) Rapid determination of neurotoxin lolitrem B in perennial ryegrass by high-performance liquid chromatography with fluorescence detection. J Chromatogr 321:217–226

Gardiner DM, Cozijnsen AJ, Wilson LM, Pedras MS, Howlett BJ (2004) The sirodesmin biosynthetic gene cluster of the plant pathogenic fungus *Leptosphaeria maculans*. Mol Microbiol 53:1307–1318

Gwinn KD, Collins-Shepard MH, Reddick BB (1991) Tissue print-immunoblot, an accurate method for the detection of *Acremonium coenophialum* in tall fescue. Phytopathology 81:747–748

Idnurm A, Howlett BJ (2003) Analysis of loss of pathogenicity mutants reveals that repeat-induced point mutations can occur in the Dothideomycete *Leptosphaeria maculans*. Fungal Genet Biol 39:31–37

Itoh Y, Johnson R, Scott B (1994) Integrative transformation of the mycotoxin-producing fungus, *Penicillium paxilli*. Curr Genet 25:508–513

Keller NP, Hohn TM (1997) Metabolic pathway gene clusters in filamentous fungi. Fungal Genet Biol 21:17–29

Kempken F, Kück U (1998) Transposons in filamentous fungi-facts and perspectives. BioEssays 20:652–659

Latch GCM, Christensen MJ (1985) Artificial infection of grasses with endophytes. Ann Appl Biol 107:17–24

Mantle PG, Weedon CM (1994) Biosynthesis and transformation of tremorgenic indole-diterpenoids by *Penicillium paxilli* and *Acremonium lolii*. Phytochemistry 36:1209–1217

McLeay LM, Munday-Finch SC, Smith BL (1999) Tremorgenic mycotoxins paxilline, penitrem and lolitrem B, the non-tremorgenic 31-epilolitrem B and electromyographic activity of the reticulum and rumen of sheep. Res Vet Sci 66:119–127

McMillan LK et al (2003) Molecular analysis of two cytochrome P450 monooxygenase genes required for paxilline biosynthesis in *Penicillium paxilli* and effects of paxilline intermediates on mammalian maxi-K ion channels. Mol Genet Genomics 270:9–23

Miles CO, Munday SC, Wilkins AL, Ede RM, Towers NR (1994) Large-scale isolation of lolitrem B and structure determination of lolitrem E. J Agric Food Chem 42:1488–1492

Möller EM, Bahnweg G, Sandermann H, Geiger HH (1992) A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues. Nucleic Acids Res 20:6115–6116

Moon CD, Tapper BA, Scott B (1999) Identification of *Epichloë* endophytes *in planta* by a microsatellite-based PCR fingerprinting assay with automated analysis. Appl Environ Microbiol 65:1268–1279

Moon CD, Scott B, Schardl CL, Christensen MJ (2000) The evolutionary origins of *Epichloë* endophytes from annual ryegrasses. Mycologia 92:1103–1118

Munday-Finch SC, Miles CO, Wilkins AL, Hawkes AD (1995) Isolation and structure elucidation of lolitrem A, a tremorgenic mycotoxin from perennial ryegrass infected with *Acremonium lolii*. J Agric Food Chem 43:1283–1288

Munday-Finch SC, Wilkins AL, Miles CO (1996) Isolation of paspaline B, an indole-diterpenoid from *Penicillium paxilli*. Phytochemisty 41:327–332

Namiki F, Matsunaga M, Okuda M, Inoue I, Nishi K, Fujita Y, Tsuge T (2001) Mutation of an arginine biosynthesis gene causes reduced pathogenicity in *Fusarium oxysporum* f. sp. *melonis*. Mol Plant Microbe Interact 14:580–584

Panaccione DG, Tapper BA, Lane GA, Davies E, Fraser K (2003) Biochemical outcome of blocking the ergot alkaloid pathway of a grass endophyte. J Agric Food Chem 51:6429–6437

Parker EJ, Scott DB (2004) Indole–diterpene biosynthesis in ascomycetous fungi. In: An Z (ed) Handbook of industrial mycology. Marcel Dekker, New York, pp 405–426

Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 85:2444–2448

Penn J, Garthwaite I, Christensen MJ, Johnson CM, Towers NR (1993) The importance of paxilline in screening for potentially tremorgenic *Acremonium* isolates. In: Easton HS (ed) Proceedings of the Second International Symposium on *Acremonium*/Grass Interactions. AgResearch Grasslands Research Centre, Palmerston North, New Zealand, pp 88–92

Proctor RH, Brown DW, Plattner RD, Desjardins AE (2003) Co-expression of 15 contiguous genes delineates a fumonisin biosynthetic gene cluster in *Gibberella moniliformis*. Fungal Genet Biol 38:237–249

Reinholz J, Paul VH (2001) Toxin-free *Neotyphodium*-isolates achieved without genetic engineering—a possible strategy to avoid "ryegrass staggers". In: Dapprich PD (ed) Fourth International *Neotyphodium*/grass Interactions Symposium. University of Paderborn-Soest, Germany, pp 261–271

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Schardl CL (1996) Interactions of grasses with endophytic *Epichlöe* species and hybrids. In: Keen N (ed) Plant–microbe interactions. Chapman and Hall, New York, pp 107–140

Schardl CL, Leuchtmann A, Tsai H-F, Collett MA, Watt DM, Scott DB (1994) Origin of a fungal symbiont of perennial ryegrass by interspecific hybridization of a mutualist with the ryegrass choke pathogen, *Epichloëtyphina*. Genetics 136:1307–1317

Schmid J, Spiering MJ, Christensen MJ (2000) Metabolic activity, distribution, and propagation of grass endophytes in planta: investigations using the GUS reporter gene system. In: White JFJ (ed) Microbial endophytes. Marcel Dekker, New York, pp 295–322

Selala MI, Laekeman GM, Loenders B, Masuka A, Herman AG, Schepens P (1991) In vitro effects of tremorgenic mycotoxins. J Nat Prod 54:207–212

Selker EU, Cambareri EB, Jensen BC, Haack KR (1987) Rearrangement of duplicated DNA in specialized cells of Neurospora. Cell 51:741–752

Siegel MR, Latch GCM, Bush LP, Fannin FF, Rowan DD, Tapper BA, Bacon CW, Johnson MC (1990) Fungal endophyte-infected grasses: alkaloid accumulation and aphid response. J Chem Ecol 16:3301–3315

Smith BL, McLeay LM, Embling PP (1997) Effects of the mycotoxins penitrem, paxilline and lolitrem B on the electromyographic activity of skeletal and gastrointestinal smooth muscle of sheep. Res Vet Sci 62:11–116

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517

Spiering MJ, Davies E, Tapper BA, Schmid J, Lane G A (2002) Simplified extraction of ergovaline and peramine for analysis of tissue distribution in endophyte-infected grass tillers. J Agric Food Chem 50:5856–5862

Springer JP, Clardy J, Wells JM, Cole RJ, Kirksey JW (1975) The structure of paxilline, a tremorgenic metabolite of *Penicillium paxilli* Bainier. Tetrahedron Lett 30:2531–2534

Toyomasu T, Nakaminami K, Toshima H, Mie T, Watanabe K, Ito H, Matsui H, Mitsuhashi W, Sassa T, Oikawa H (2004) Cloning of a gene cluster responsible for the biosynthesis of diterpene aphidicolin, a specific inhibitor of DNA polymerase alpha. Biosci Biotechnol Biochem 68:146–152

Trinci APJ (1978) The duplication cycle and branching in fungi. In: Trinci APJ (ed) Symposium of the British Mycological Society, Queen Elizabeth College, London, pp 319–357

Tudzynski B, Hölter K (1998) Gibberellin biosynthetic pathway in *Gibberella fujikuroi*: evidence for a gene cluster. Fungal Genet Biol 25:157–170

Vollmer SJ, Yanofsky C (1986) Efficient cloning of genes of *Neurospora crassa*. Proc Natl Acad Sci USA 83:4869–4873

Wu R, Hirai A, Mundy J, Nelson R, Rodriguez R (1991) Guidelines for nomenclature of cloned genes or DNA fragments in rice. Rice Genet Newslett 8:51–53

Yoder OC (1988) *Cochliobolus heterostrophus*, cause of southern corn leaf blight. Adv Plant Pathol 6:93–112

Young C, Itoh Y, Johnson R, Garthwaite I, Miles CO, Munday-Finch SC, Scott B (1998) Paxilline-negative mutants of *Penicillium paxilli* generated by heterologous and homologous plasmid integration. Curr Genet 33:368–377

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

Zhang S, Monahan BJ, Tkacz JS, Scott B (2004) An indole--diterpene gene cluster from *Aspergillus flavus*. Appl Environ Microbiol 70:6875--6883

Table 1 Sequence and bioinformatics analysis of *ggsA* and *ltm* genes

Gene[a]	Proposed function of product	Size (aa)	MW (kDa)	Intron number	Homologue in *pax* gene cluster	Degree of product identity (%)
NIggsA	GGPP synthase	466	51.9	0		
EfggsA	GGPP synthase	465	51.8	0		
EtggsA	GGPP synthase	464	51.6	0		
NIltmG	GGPP synthase	334	37.9	2	*paxG*	52.6
NIltmM	Monooxygenase	472	52.5	3	*paxM*	41
NIltmK	P450 monooxygenase	533	60.9	7	*paxP/Q*	31.3

[a] The analysis of the *E. festucae ltmG*, *ltmM* and *ltmK* sequences is identical to *N. lolii* with the exception of one amino acid difference in *E. festucae* LtmM.

Table 2 Fungal biomass and alkaloid production

Strain	Genotype[a]	Number of plants examined	Endophyte biomass (%)[b]	Lol (ppm)[c]	Erg (ppm)[c]	Per (ppm)[c]
PN2303	*ΔltmM*	5	0.4–1.5	0	0.4–1.3	30–40
PN2296	*ΔltmM*	5	0.4–1.5	0	0.1–2.1	15–47
PN2301	*ΔltmGM*	4	0.4–1.9	0	0.7–3.3	24–41
PN2294	Ectopic	5	0.7–1.5	4.4–16.7	0.5–1.2	21–55
FI1	wt	4	1.0–1.7	6.2–12.8	0.8–1.5	31–66
Endophyte free	NA	3	NA	0	0	0

[a] *ΔltmM*, ltmM insertional replacement; (*Δ ltmGM*, extended deletion mutant; wt, wild-type; NA, not applicable. [b] Endophyte biomass was determined from the ratio of endophyte to plant DNA measured by real-time quantitative PCR. [c] There was no significant ($P<0.05$) differences among samples as determined by analysis of variance. Lol, lolitrem; Erg, ergovaline; Per, peramine.

Table 3 Complementation of the *E. festucae ltmM* deletion mutant

Strain	Complementation plasmid[a]	Number of plants assayed	Lolitrem (ppm)[b]
PN2338	pCY40	2	2.6
PN2339	pCY40	3	2.9
PN2340	pCY40	2	3.2
PN2341	pCY40	2	0.6
PN2350	pCY41	2	0
PN2351	pCY41	2	0
PN2352	pCY41	2	1.4
PN2346	pII99	2	0
PN2347	pII99	2	0
PN2303	NA	2	0
FII	NA	2	3.9
Endophyte free	NA	2	0

[a] Complementations were carried out in the PN2303 background. pCY40 contains the *ltmM* gene on a 7-kb *Xho*I fragment. pCY41 contains the *paxM* gene on a 3.7-kb *Bgl*II fragment. [b] Average value.