Characterisation of novel secondary metabolism genes in plant-endophytic *Epichloë* fungi

A thesis presented in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in

Genetics

at Massey University, Manawatu,

New Zealand

Daniel Berry

2016

ABSTRACT

Plant-endophytic fungi of the genus *Epichloë* are symbionts of cool season grasses. *Epichloë* infections are generally asymptomatic, with endophyte and host growth synchronised to form an interconnected hyphal network in the intercellular spaces separating host cells. In return for the sustenance received from the host apoplast, *Epichloë* spp. produce a range of bioprotective secondary metabolites (SM) that deter both vertebrate and invertebrate herbivores. Peramine is an *Epichloë*-derived insect-deterring SM produced by the two-module non-ribosomal peptide synthetase (NRPS) PerA, encoded by the gene *perA*. The *perA* gene is widespread across *Epichloë*, but peramine production is much more discontinuous. Through an *Epichloë*-wide survey it is shown that the peramine chemotype of *Epichloë* isolates can be accurately predicted through a combination of PCR and sequencing methods. Furthermore, the protein encoded by the widespread peramine-negative *perA-ΔR* allele is analysed in detail. The *perA-ΔR* allele has a transposable element inserted into the 3' end of the gene, and was previously assumed to be non-functional. Evidence for ongoing *perA-ΔR* functionality is presented, with the PerA-ΔR protein shown to exhibit altered substrate specificity relative to PerA. A PerA/PerA-ΔR domain swap is shown to restore peramine production to PerA-ΔR, providing insight into the functional changes that separate this protein from PerA.

Another major project investigated the role SM genes may play in *Epichloë* sexual development. Many *Epichloë* spp. are transmitted vertically through infection of the host seed. However, sexual *Epichloë* spp. may also undergo a reproductive cycle that involves formation of pre-sexual stromata structures over developing host inflorescences. This sexual cycle culminates in horizontal transfer of the endophyte to endophyte-free host plants. Deletion of the symbiotically regulated endophyte gene *irlA* induced a symbiosis-defective phenotype *in planta*, and this observation led to the identification of a novel *Epichloë* SM cluster. Synteny analysis and comparison to previously characterised fungal SM gene clusters defines a five-gene cluster centred on *irlA* that is shown to be upregulated in stromata. A model is

proposed whereby this SM cluster controls initiation of the proliferative hyphal growth associated with stromata formation.

ACKNOWLEDGEMENTS

Firstly, a big thank you to my primary supervisor Prof. Barry Scott, whose outstanding guidance throughout both my MSc and PhD has enabled me to reach this point. Thank you for the unceasing enthusiasm and support you bought to my projects. Thanks also to my co-supervisors A. Prof. Carolyn and Dr Paul Dijkwel, who have both been there throughout this long process to provide the guidance and support required.

I would like to express my gratitude to Dr Katrin Grage for her assistance with the PerA-ΔR project. Without Katrin's help it would not have been possible to get this project to the point it is at today. Thanks also to Dr. Wade Mace of AgResearch, whose chemical expertise continues to be of great assistance to our collaboration on the PerA-ΔR project. Thanks also Prof. Murray Cox for his assistance in assembling the *P. paxilli* genome, and to Arvina and all the members of Scottbase lab that have helped me over the years. Stepping into the wider Massey University community, the services provided by the MMIC and the Massey Genome Service were outstanding and invaluable for completing these projects.

The projects presented within this thesis would not have been possible without the collaborations formed with members of the wider research community. In particular, I would like to acknowledge the immensely helpful genomics resources and insight provided by Prof. Chris Schardl at the University of Kentucky, who also kindly gave access to unpublished genomes and RNAseq data to aid these projects. I also thank Dr. Richard Johnson from AgResearch for providing sequence from the AR1 and AR37 genomes; Devish Singh (Barenbug) and A. Prof Shaun Bush (USDA) for providing access to their field trials to obtain stromata samples; Prof. Geoff Jameson for his insight into IrlA function and A. Prof David Ackerley for his insights on NRPS functionality. Thanks also to my biology teacher Mike Seawright, who went above and beyond the call of duty and undoubtedly influenced where I am today.

Grateful acknowledgments are made to the financial support provided by the Massey University Doctoral Scholarship, the BioProtection Research Centre and the RSNZ Marsden fund. Thanks also to the Samuel Roberts Noble Foundation for

hosting me on several occasions, those visits were hugely helpful for enabling the research presented here.

Finally, thanks to my parents for the unconditional love and support they have provided throughout my years of studies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

COMMON ABBREVIATIONS

10AA	The 10 amino acid A-domain substrate specificity-conferring code
4'PP	4'-phosphopantetheine
4'PPTase	4'-Phosphopantetheinyl transferase
A-domain	NRPS AMPylation domain
aa	amino acid(s)
ABI	Applied Biosystems Inc.
AFR	Adenylate-forming reductase
Amp	Ampicillin
AMP	Adenosine monophosphate
Amp^R	Ampicillin resistance
ATP	Adenosine triphosphate
BBE	Berberine bridge enzyme
BLAST	Basic local alignment search tool
BLASTn	BLAST search against a nucleotide sequence database with a nucleotide sequence query
BLASTp	BLAST search against a protein sequence database with a protein sequence query
bp	Base pairs
C-domain	NRPS condensation domain
cDNA	Complementary DNA
CDS	Coding sequence
CDYE	Czapek Dox yeast extract
CLSM	Confocal laser-scanning (fluorescence) microscopy
CRISPR	Clustered regularly-interspaced short palindromic repeats
DEPC	Diethylpyrocarbonate
DIG	Digoxigenin
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	deoxyribose nucleoside triphosphate

E^-	Not infected by endophyte
EAS	Ergot alkaloid synthesis
EDTA	Ethylene diamine tetra-acetic acid
eGFP	Enhanced green fluorescent protein
FAD	Flavin adenine dinucleotide
g	Acceleration due to gravity on Earth (9.8 m s^2)
gDNA	Genomic DNA
Gen	Geneticin
Gen^R	Geneticin resistance
GFP	Green fluorescent protein
h	Hour(s)
hph	Hygromycin resistance-conferring gene
Hyg	Hygromycin
Hyg^R	Hygromycin resistance
IDT	Indole-diterpene
IFR	Isoflavone reductase
IRL	Isoflavone reductase-like
kb	kilo base pairs
LB	Lysogeny broth
LCMS	Liquid chromatography-mass spectrometry
LOL	Loline
M	Moles per liter
M-domain	NRPS methylation or methyltransferase domain
min	Minute(s)
mRNA	Messenger RNA
MS or MS^1	Mass spectrometry
MS^2	Tandem mass spectrometry
NAD	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
NCM	Nitrocellulose membrane
nptII	Geneticin resistance-conferring gene

NRP	Non-ribosomal peptide
NRPS	Non-ribosomal peptide synthetase
P5C	1-pyrroline-5-carboxylate
PAMP	Pathogen-associated molecular pattern
PAX	Paxilline
PCP	Peptide carrier protein
PCR	Polymerase chain reaction
PD	Potato dextrose
PEG	Polyethylene glycol
PER	Peramine
PIP	**P**inoresinol-lariciresinol reductase/**I**soflavone reductase/ **P**henylcoumaran benzylic either reductase family proteins
PKS	Polyketide synthetase
PLP	Pyridoxal phosphate
PP_i	Phyrophosphate
qPCR	Quantitative polymerase chain reaction
R-domain	Reductase domain
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
RNase	Ribonuclease
RPM	Revolution per minute
RT-PCR	Reverse transcription polymerase chain reaction
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
s	Second(s)
SAM	S-adenosylmethionine
SDS	Sodium dodecyl sulfate
SM	Secondary metabolism OR secondary metabolite
SNP	Single nucleotide polymorphism
T-domain	NRPS thiolation domain (aka PCP domain)
TBE	Tris-boric acid-EDTA

tBLASTn	BLAST search against a translated nucleotide sequence database with a protein sequence query
TE	Trace elements (or thioesterase; context dependent)
TE-domain	NRPS type I thioesterase domain
TEII	Type II thioestease
TF	Transcription factor
TM	Trans-membrane
TSP	Trans-species polymorphism
v/v	Volume:volume ratio
w/v	Weight:volume ratio
WGA	Wheat germ agglutinin
wt	Wild type

1 Introduction

1.1 *Epichloë* spp., endophytes of cool-season grasses

Fungi of the genus *Epichloë* are endophytes of cool-season grasses from the subfamily Poöideae (Clay and Schardl, 2002). These endophytes reside within the intercellular spaces of host aerial tissues, where they obtain sustenance from the host apoplastic fluid. Infection of above-ground tissues originates from endophytic hyphae within the shoot apical meristem, with these hyphae infecting developing leaf primordia and axillary buds by tip growth (Christensen and Voisey, 2007). Intercalary extension of hyphae within the leaf expansion zone synchronises fungal growth with the rapidly enlarging plant cells, to which the hyphae are strongly attached (Christensen et al., 2008). Growth of these endophytic hyphae ceases when leaf growth stops (Christensen and Voisey, 2007), but metabolic activity is maintained (Tan et al., 2001). This results in a systemic network of hyphae within mature grass pseudostems and blades, typically consisting of singular hyphae contained within intercellular spaces along the leaf axis, with occasional lateral branches that fuse to parallel hyphae in nearby intercellular spaces to establish and maintain an interconnected hyphal network (Figure 1.1). Many *Epichloë* spp. also produce epiphytic hyphae that grow on the surface of the host cuticle (Christensen and Voisey, 2007). These epiphytic hyphae derive from endophytic hyphae that have penetrated through the host cuticle by formation of an expressorium, an appresorium-like structure (Becker et al., 2016). These epiphytic hyphae often produce conidia that have the potential to mediate asexual horizontal transfer to endophyte-free hosts (Tadych et al., 2012). Interestingly, microscopic analysis of endophyte-infected plants suggests that chitin in the cell wall of *Epichloë* hyphae is masked or removed in endophytic hyphae, presumably to avoid a PAMP triggered host defence response (Becker et al., 2015; Dupont et al., 2015), but the chitin is clearly visible for epiphytic hyphae (Becker et al., 2016).

While endophytic growth largely ceases when plant growth stops, the hyphae themselves remain metabolically active, with *Epichloë* endophytes continuing to produce a number of different secondary metabolites (Schardl et al., 2012) and proteins (Ambrose et al., 2014) that protect the host against a wide variety of

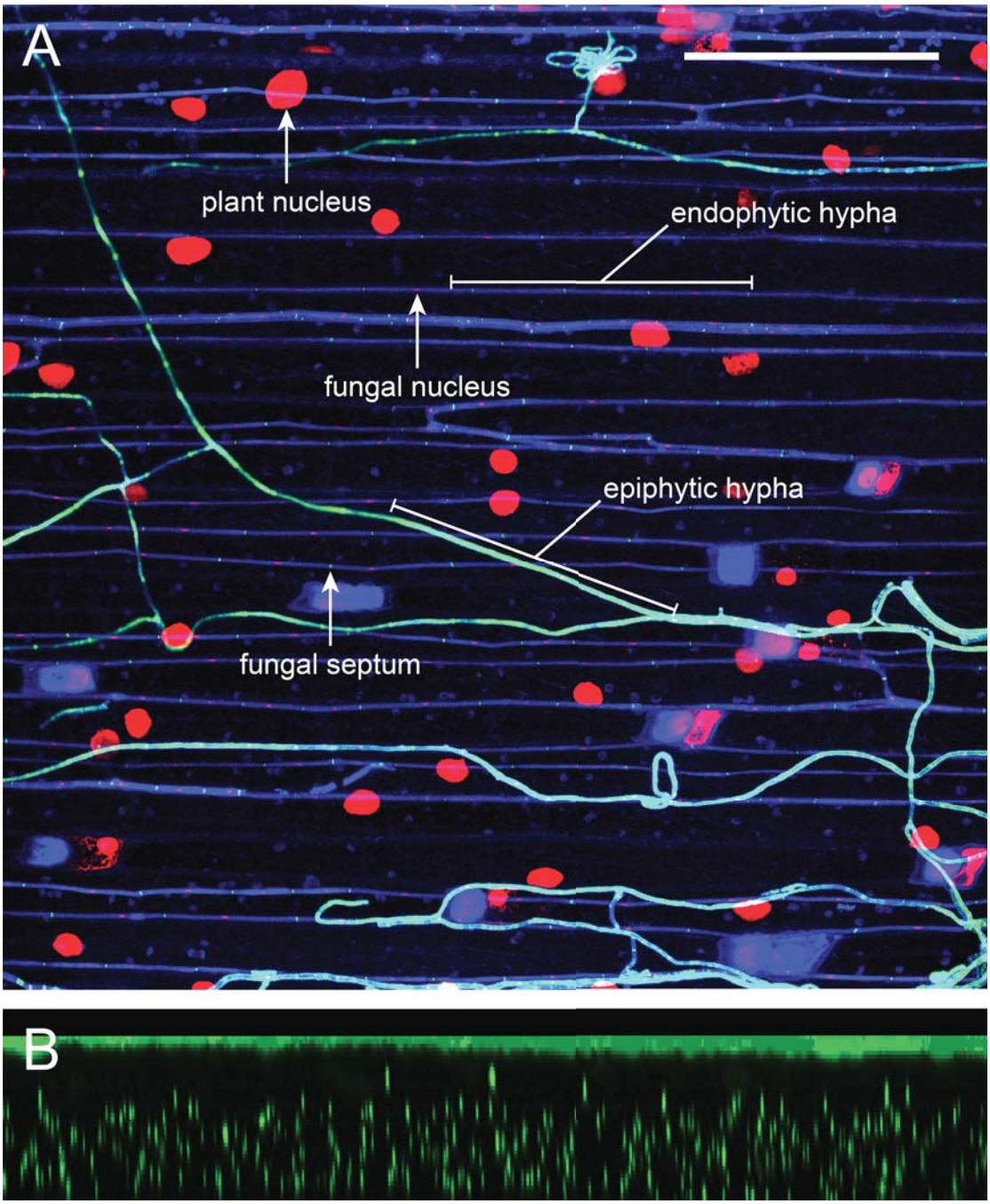


Figure 1.1 ***E. festucae*** **hyphae growing** ***in planta*****.** (A) confocal laser-scanning fluorescent microscopy image of *E. festucae* hyphae growing in *Lolium perenne* with fungal hyphae stained with aniline blue (pseudocolour blue), chitin in fungal septa/epiphytic cell walls stained with WGA-AlexaFluor® 488 (pseudocolor green) and nuclei stained with propidium iodide (pseudocolour red). Image shows z-stack of images taken 0.46 µm apart through an approx. 16 µm thick section taken from surface of leaf sheath downwards. Scale bar = 100 µm. (B) 3D projection of top image showing side-on view with green stain only, demonstrating that the fully green hyphae are on the pseudostem surface only (vertical scale increased 400% for clarity).

herbivores. *Epichloë* spp. have also been shown to increase host tolerance to abiotic stresses such as drought (Arachevaleta et al., 1989; Hahn et al., 2008), but the mechanisms underlying these benefits are less clear. During host reproduction many *Epichloë* spp. can be vertically transmitted to a new generation of plants though the host seed by infection of developing ovules (Schardl, 2010). Alternatively, the sexual cycle of *Epichloë* spp. culminates in horizontal transmission of the endophyte by a *Botanophila* sp. fly (Bultman and Leuchtmann, 2008; Tanaka et al., 2012), and is synchronised with host flowering, involving reproductive sterilisation of the infected tiller (Schardl, 2010). While *Epichloë* spp. occur naturally only in association with their grass hosts, many are easily isolated from infected plant material, and can be cultured *ex planta* indefinitely on relatively simple media (Bacon, 1988).

1.1.1 The symbiotic spectrum

Epichloë species and strains exist within a symbiotic spectrum that ranges from highly mutualistic to largely antagonistic (Schardl and Leuchtmann, 2005). The position of any strain within this spectrum is primarily determined by the method of transmission and types of bio-protective molecules produced. Distribution of the secondary metabolism (SM) genes required for production of most bio-protective molecules is highly discontinuous across *Epichloë*, resulting in wildly differing chemotypes between *Epichloë* strains (Clay and Schardl, 2002; Schardl et al., 2012; Schardl et al., 2013a; Young et al., 2009). Specific *Epichloë* species and strains may thus confer varying levels of protection against herbivores to their host (Schardl et al., 2012; Schardl et al., 2013b). This protection represents a large fraction of the potential benefits that the host can gain from these associations (Schardl et al., 2012).

Because the sexual cycle of *Epichloë* spp. involves reproductive sterilisation of the infected host tiller, the frequency at which an *Epichloë* isolate initiates sexual development is another important determinant for where it is located within the symbiotic spectrum (Schardl and Leuchtmann, 2005). *Epichloë* spp. are classed as type I, type II or type III based on how the isolate reproduces and disseminates (Clay and Schardl, 2002). Type I endophytes only disseminate after sexual reproduction and host reproduction in plants infected with type I endophytes is generally severely

impacted, with inflorescences aborted on most or all infected tillers (Schardl and Leuchtmann, 2005). In contrast, type III endophytes never undergo sexual reproduction and are only disseminated through the host seed, allowing host inflorescences to develop fully, whereas type II endophytes are capable of both methods of dissemination (Clay and Schardl, 2002; Schardl and Leuchtmann, 2005).

1.1.2 *Epichloë* sexual reproduction

The sexual cycle of *Epichloë* spp. is synchronised with host flowering, and starts with formation of a dense mass of interwoven epiphytic hyphae (pseudoparenchyma) called a stroma (pl. stromata) that envelops the developing host inflorescence (Figure 1.2; Clay and Schardl, 2002). These unfertilised stromata are coloured white and prevent host flowering, causing reproductive sterilisation of affected tillers. The frequency at which type II *Epichloë* endophytes form stromata shows considerable genotypic variation (Schardl and Leuchtmann, 2005); however, the molecular mechanisms underpinning this variation are not known. Stromata produce spermatia that are transferred to other stromata by *Botanophila* flies, which in turn lay their eggs on *Epichloë* stromata so that their larvae may feed off them once hatched (Bultman and Leuchtmann, 2008). *Epichloë* spp. are heterothallic, meaning that exchange of genetic material between opposite mating types is required for sexual reproduction (Clay and Schardl, 2002). If the donor and recipient stromata of these spermatia are of opposite mating types, the recipient stroma becomes fertilised, resulting in thickening of the stroma, production of a yellow-orange pigment and production of perithecia (Figure 1.2). These perithecia eject meiotically-derived ascospores that enable horizontal transfer to uninfected plants (Schardl, 2010). This horizontal transmission has been shown to occur through infection of seed in nominally uninfected plants, presumably through invasion of the ovule while the grass is flowering, resulting in systemic infection of the germinated seedlings (Chung and Schardl, 1997). Additionally, conidia and ascospore-vectored horizontal transmission by direct infection of vegetative tissues in uninfected seedlings and established plants has also been observed (Brem and Leuchtmann, 1999; Leyronas and Raynal, 2008; Tadych et al., 2012).

Figure 1.2 ***E. typhina*** **stromata on orchardgrass.** (Top) Unfertilised *E. typhina* stroma with *Botanophila* sp. fly on aborted *Dactylis glomerata* (orchardgrass) inflorescence. (Bottom) Fertilised *E. typhina* stromata on *D. glomerata*.

1.1.3 Epichloë phylogeny, host specificity and hybridisation

There have been 10 sexual *Epichloë* spp. described to date, along with an ever-increasing number of asexual species (Leuchtmann et al., 2014). A cladogram approximating *Epichloë* ancestry is shown in Figure 1.3. These *Epichloë* spp. are restricted to hosts from the grass subfamily Poöideae, within which a specific *Epichloë* sp. typically infects hosts from a specific genus or group of closely related genera (Schardl, 2010). In recent years the amount of DNA sequence data available

from *Epichloë* isolates has exploded, with the genomes of 18 strains from 14 different *Epichloë* species and subspecies currently publicly available along with a variety of genomes from the wider Clavicipitaceae (www.endophyte.uky.edu; Schardl et al., 2013b), providing an invaluable resource for molecular biology in these organisms.

Asexual *Epichloë* spp. were formerly classified in the form genus *Neotyphodium*, but were recently consolidated into *Epichloë* to better reflect their evolutionary origins (Leuchtmann et al., 2014). While some asexual *Epichloë* spp. are derivatives of a single sexual *Epichloë* spp., most are the result of inter-specific hybridisation between two or more usually sexual *Epichloë* spp. (Selosse and Schardl, 2007). These hybrids generally contain one copy of each gene from each donor strain, excepting where gene deletions have occurred, necessitating multiple taxonomic classifications for each hybrid isolate (e.g. LpTG-2 is a hybrid of *E. festucae* × *E. typhina*). Asexual *Epichloë* spp. are type III endophytes that are only transmitted vertically through the host seed and generally do not form stromata, though there have been occasional reports of this occurring in certain asexual *Epichloë* spp., though no perithecia were formed (Ji et al., 2009; Schardl, 2010).

1.1.4 *Epichloë* secondary metabolism

Epichloë spp. as a whole produce a variety of different bio-protective secondary metabolites, also commonly referred to as natural products (Schardl et al., 2012). The most well studied examples of these are the alkaloids synthesised by the indole-diterpene (*IDT*), ergot alkaloid (*EAS*), loline alkaloid (*LOL*) and peramine (*PER*) biosynthetic pathways (Fleetwood et al., 2007; Haarmann et al., 2005; Pan et al., 2014a; Panaccione et al., 2001; Schardl et al., 2013a; Spiering et al., 2005; Tanaka et al., 2005; Tsai et al., 1995; Wang et al., 2004; Young et al., 2005; Young et al., 2006). These pathways often require the proteins encoded by a large number of different genes to enable product synthesis. For example, production of the indole-diterpene lolitrem B, which is a potent tremorgen in vertebrates, requires the proteins encoded by at least 10 separate *IDT* genes (Saikia et al., 2012). Synthesis of the mycotoxic ergot alkaloid ergovaline and the anti-insert loline alkaloid *N*-formylloline require similarly large numbers of genes (Fleetwood et al., 2007; Pan et

al., 2014a; Pan et al., 2014b; Panaccione et al., 2001; Schardl et al., 2013a; Spiering et al., 2008; Spiering et al., 2005; Spiering et al., 2002; Wang et al., 2004); however, synthesis of peramine is proposed to only require a single gene, *perA* (Tanaka et al., 2005). Peramine is a pyrrolopyrazine-containing molecule that provides feeding deterrence against some invertebrate herbivores, including the agriculturally relevant pest *Listronous bonariensis* (Argentine stem weevil; Rowan et al., 1990).

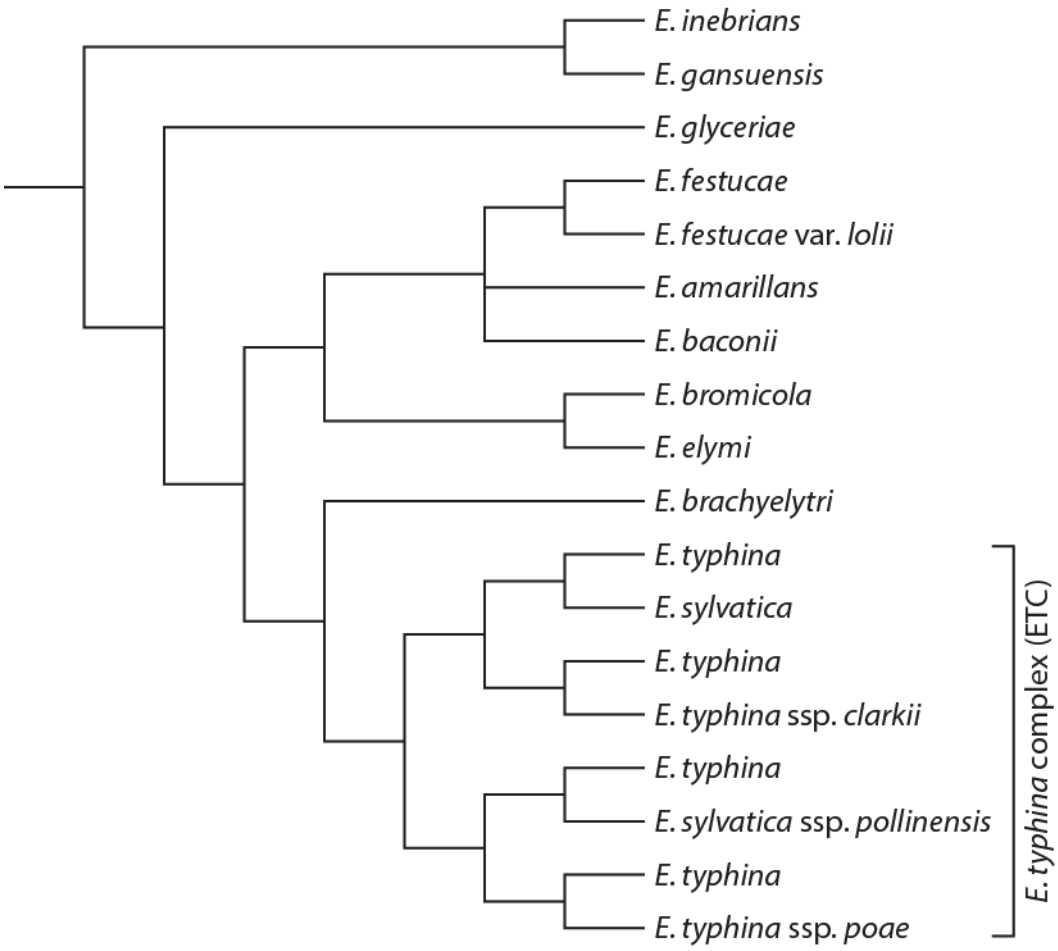


Figure 1.3 Cladogram approximating the taxonomy of representative non-hybrid *Epichloë* spp. Cladogram is a simplification of the results presented by Leuchtmann et al. (2014).

The genes for the *IDT*, *EAS* and *LOL* pathways are organised into contiguous clusters within their *Epichloë* host genomes (Fleetwood et al., 2007; Spiering et al., 2005; Young et al., 2006). The *IDT* and *EAS* clusters are located proximate to telomeres, and the *IDT*, *EAS*, *LOL* and occasionally *PER* loci were associated with a

large number of repetitive sequence relics of transposable elements (Fleetwood et al., 2011; Schardl et al., 2013a; Young et al., 2005; Young et al., 2006). The dynamic chromosomal environment that these clusters reside in is proposed to enable rapid evolutionary innovation within these SM pathways (Fleetwood et al., 2011; Schardl et al., 2013a; Young et al., 2005; Young et al., 2006). Gene loss is common within these regions, and is likely the cause of the highly discontinuous distribution of *IDT*, *EAS* and *LOL* genes across *Epichloë* (Schardl et al., 2013a). As a result, many *Epichloë* spp. contain subsets of the *IDT*, *EAS* or *LOL* genes that enable production of what are pathway intermediates in the most complex examples of these biosynthetic pathways as a final product (Schardl et al., 2012; Schardl et al., 2013a; Schardl et al., 2013b; Young et al., 2015). Alternatively, gene loss events have resulted in the deletion of entire SM clusters (Schardl et al., 2013a). Unlike the *IDT* and *EAS* gene clusters, *perA* is not located near a telomere (Schardl et al., 2013a). This relatively stable genomic locus might explain why *perA* is found in almost all *Epichloë* spp. However, peramine production is also discontinuous across *Epichloë*, indicating a prevalence of *perA* null alleles (Clay and Schardl, 2002; Schardl et al., 2012). One such *perA* allele incapable of producing peramine is *perA-ΔR*, which contains a deletion in the 3' end of the gene associated with the insertion of a miniature inverted-repeat transposable element (Fleetwood et al., 2011; Schardl et al., 2013a). Nevertheless, peramine is still the most widespread of the identified bio-protective alkaloids across *Epichloë* (Clay and Schardl, 2002; Schardl et al., 2012).

Production of any given *Epichloë* secondary metabolite is notably more common in asexual hybrid *Epichloë* spp. (Schardl et al., 2013b). This can be partially explained by the presence of SM genes derived from two or even three different *Epichloë* spp. in these isolates. However, some hybrid isolates contain SM genes derived from sexual *Epichloë* spp. for which these SM clusters have not been described in surveys of extant strains (Schardl et al., 2013b). This abnormally large representation of apparently rare genotypes was proposed to be consistent with the hypothesis that asexuality in *Epichloë* spp. selects for mutualism. *Epichloë* spp. that produced these secondary metabolites would have a selective advantage to the host over non-producers (Ewald, 1987; Schardl, 2010; Schardl et al., 2013b). This hypothesis is also consistent with the observation that bio-protective SM genes are

less common in the relatively pathogenic type I endophytes than in type II and type III non-hybrid endophytes (Schardl et al., 2013a).

1.1.4.1 Comparison of SM in *Epichloë* to other fungi

Clustering of genes belonging to the same metabolic pathway into contiguous units is commonly observed in fungi, and is particularly prevalent for SM genes (Keller and Hohn, 1997; Keller et al., 2005). *Epichloë* spp. are no exception to this trend, with all *EAS*, *LTM* and *LOL* genes organised into pathway-specific clusters (Fleetwood et al., 2007; Schardl et al., 2013a; Spiering et al., 2005; Young et al., 2006). Gene clustering enables the acquisition, loss or transfer of genes belonging to the same pathway to occur much more efficiently, and is thought to be a major causative mechanism for the huge diversity of secondary metabolites observed in fungi (Wisecaver and Rokas, 2015; Wisecaver et al., 2014). Gene clustering is also proposed to facilitate chromatin-level regulation of genes belonging to the same metabolic pathway (Keller and Hohn, 1997). Indeed, such chromatin-level regulatory mechanisms have been recently documented for SM genes of fungi in general (Gacek and Strauss, 2012), as well as for SM genes in *E. festucae* (Chujo and Scott, 2014). The *Epichloë IDT*, *EAS*, *LOL* and *PER* pathway genes are all symbiotically regulated, with significant gene expression only observed *in planta* (Fleetwood et al., 2007; Spiering et al., 2005; Tanaka et al., 2005; Young et al., 2006). Abolishment of SM-specific heterochromatic marks was shown to partially de-repress *E. festucae* SM genes in culture, but was not sufficient to induce the levels of expression observed *in planta* (Chujo and Scott, 2014). This indicates that a combination of regulatory mechanisms, such as *in planta*-specific transcription factor(s), are responsible for establishing the symbiotic regulation observed for *Epichloë* SM genes. Regulation of gene expression for many fungal SM gene clusters is established by a pathway-specific transcription factor encoded within the gene cluster it controls (Keller et al., 2005). This is not the case in *Epichloë* spp., with all SM clusters identified to date lacking a pathway-specific transcription factor (Fleetwood et al., 2007; Schardl et al., 2013a; Spiering et al., 2005; Tanaka et al., 2005; Young et al., 2006).

The plant-specific expression of *Epichloë* SM genes complicates characterisation of the encoded proteins, as *Epichloë* spp. must be inoculated into

seedlings, grown for several weeks and then tested for host infection before metabolite analyses can be performed (Fleetwood et al., 2007; Pan et al., 2014a; Panaccione et al., 2001; Tanaka et al., 2005; Young et al., 2005). These analyses then also have the added complexity of the plant metabolic background, which confounds analysis when attempting to identify a novel secondary metabolite. Due to this added complexity, analysis of the *IDT* genes in *Epichloë* spp. was preceded by analysis of *Penicillium paxilli IDT* genes (aka *PAX* cluster), which produce the indole-diterpene paxilline (McMillan et al., 2003; Scott et al., 2013; Young et al., 1998; Young et al., 2001). Compared to *Epichloë* spp. *P. paxilli* is faster growing, sporulates much more readily and produces paxilline in culture. Cloning the *IDT* genes from *P. paxilli* enabled the identification and characterisation of the *IDT* gene cluster in *E. festucae* (Young et al., 2005; Young et al., 2006). These experiments included complementation of specific *P. paxilli IDT* gene deletion mutants by heterologous expression of *E. festucae IDT* genes to test for functional equivalency (Saikia et al., 2012; Young et al., 2006).

While deletion or post-transcriptional silencing of SM genes from the *IDT*, *EAS*, *LOL* and *PER* loci eliminates production of their respective bio-protective metabolites, there is no effect on the symbiotic interaction phenotype (Fleetwood et al., 2007; Pan et al., 2014a; Spiering et al., 2005; Tanaka et al., 2005; Young et al., 2005). The dispensable nature of these genes for growth under normal laboratory conditions is typical for fungal SM genes, which are often expressed only under specific environmental conditions (Keller et al., 2005). However, this should not be taken to mean that secondary metabolites are not important; secondary metabolites have been shown to be necessary for a variety of essential functions, such as pathogenicity (Scharf et al., 2014), synthesis of intracellular siderophores (Oide et al., 2007) and sclerotia formation (Forseth et al., 2013). A deletion mutant of the SM gene *sidN* in *E. festucae*, which produces an extracellular siderophore, has also been shown to cause a symbiosis-defective phenotype in association with *L. perenne* (Johnson et al., 2013b).

The *IDT* and *EAS* gene clusters are widespread throughout the Clavicipitaceae, though the products produced by these pathways vary considerably between genera and species as the result of divergence in pathway gene

composition and sequence (Schardl et al., 2013a; Young et al., 2015). *LOL* genes have only been identified in the genomes of *Epichloë* spp. and *Atkinsonella hypoxylon* to date (Pan et al., 2014a; Schardl et al., 2013a), though lolines have also been detected in material from plants beyond the known host range of *Epichloë* and *Atkinsonella* spp. (Schardl et al., 2007; Tofern et al., 1999). Peramine production has not been reported beyond *Epichloë* spp., although a *perA* pseudogene has been identified in the genome of the closely related fungus *Periglandula ipomoeae*, which is an *Ipomoeae asarifolia* (morning glory) symbiont (Schardl et al., 2013a). This is the only *perA* homolog identified outside of *Epichloë* to date. Narrow taxonomic distribution of specific secondary metabolites is typical in fungi (Keller et al., 2005). Nevertheless, secondary metabolites can generally be organised into one of several classes, with metabolites from each class being synthesised by shared core biosynthetic machinery, while innovation in the mid or late stages of these pathways provides the observed metabolic diversity (Keller et al., 2005; Schardl et al., 2013b; Young et al., 2015). Two particularly diverse secondary metabolite classes are the polyketides and non-ribosomal peptides, synthesised by polyketide synthetases (PKS) and non-ribosomal peptide synthetases (NRPS), respectively (Keller et al., 2005). NRPSs and PKSs are modular proteins often referred to as megasynthases due to their enormous size. Each module in these megasynthases incorporates a single monomeric substrate into a growing polyketide or polypeptide chain (Cox and Simpson, 2009; Finking and Marahiel, 2004). The modular nature of these proteins, combined with extreme substrate diversity and the activity of modifying domains, results in the enormous diversity of polyketides and non-ribosomal peptides observed in fungi (Cox and Simpson, 2009; Eisfeld, 2009).

1.2 Non-ribosomal peptide synthetases

Non-ribosomal peptide synthetases are large multi-modular enzymes found in bacteria and fungi that produce small peptide products independently from the ribosomal machinery (Finking and Marahiel, 2004). A typical NRPS contains an initiation module at its N-terminus, which begins peptide synthesis, followed by a number of elongation modules. Every module is responsible for the incorporation of a single specific amino acid, and the ordering of modules within the NRPS explicitly

defines the amino acid sequence of the final peptide product (Marahiel et al., 1997). Each module contains an AMPylation domain (A-domain; previously known as adenylation domains) that selects and activates a specific substrate, followed by a thiolation domain (T-domain; also known as a peptidyl carrier protein or PCP) that anchors the activated substrate via a covalent thioester bond (Finking and Marahiel, 2004). In addition to these two domains, every elongation module contains a condensation domain (C-domain) at its N-terminus that catalyses peptide bond formation between the substrate moieties bound to the current and immediately preceding modules (Finking and Marahiel, 2004). Any NRPS module may also contain one or more modification domains, such as methylation, oxidation or epimerisation domains, which alter the bound substrate of that module prior to peptide bond formation (Hur et al., 2012). The final module in any NRPS typically contains a termination domain located at its C-terminus that is responsible for release of the final peptide product from the NRPS (Du and Lou, 2010).

1.2.1 AMPylation domains

NRPSs are able to incorporate a large variety of proteogenic and non-proteogenic amino acids, including D-amino acids and occasionally imino, carboxy or hydroxyl acids, into their peptide products (Caboche et al., 2010; von Döhren, 2004). The ~550 aa long A-domains are responsible for the initial substrate specificity of each NRPS module, as well as substrate activation. This activation consists of AMPylation of the substrate carboxyl group in the reaction ATP + carboxyl group → acyl-AMP + PP_i (Conti et al., 1997). The specificity of any given A-domain is primarily defined by the identity of the ten amino acid residues lining the substrate-binding pocket (Conti et al., 1997), though other residues within the A-domain can also contribute (Conti et al., 1997; Lee et al., 2010; Rausch et al., 2005). The high conservation of protein secondary/tertiary structure and core sequence motifs across A-domains (Marahiel et al., 1997) enables highly accurate prediction of the binding pocket residues, allowing a specificity-defining codon or '10AA code' to be generated for each A-domain (Challis et al., 2000; Stachelhaus et al., 1999). Dissection of this 10AA code from a large number of A-domains with known substrate specificities (Caboche et al., 2008) has enabled computational substrate

prediction for uncharacterised A-domains with reasonable accuracy (Prieto et al., 2012; Roettig et al., 2011; Weber et al., 2015). However, a dearth of fungal A-domains with characterised substrates means these computational predictions are less accurate for fungi-derived NRPS sequences compared to bacterial sequences (Eisfeld, 2009; Lee et al., 2010). Nevertheless, for amino acid-binding A-domains the first residue of this 10AA code is invariably aspartic acid, as the carboxyl side chain of this residue stabilises the substrate amino acid via hydrogen bond formation with the substrate α-amino group (Conti et al., 1997). NRPS A-domains are not restricted to binding amino acid substrates, and the absence of an aspartic acid residue at this location is a strong indication of a non-amino acid substrate (Schneider et al., 2003), such as an α-keto acid (Schneider et al., 2008) or an α-hydroxy acid (Donzelli et al., 2012).

1.2.2 Thiolation domains

The 80-100 aa long T-domains are the structural anchors to which the amino acid substrates and resulting peptides of any NRPS are attached (Hur et al., 2012). A single T-domain is an essential component of every NRPS module, within which it is usually fused to the C-terminus of an upstream A-domain and the N-terminus of a downstream condensation or termination domain (Finking and Marahiel, 2004). T-domains require post-translational modification by an external 4'-phosphopantetheinyl transferase (4'PPTase), which attaches a Coenzyme A (CoA)-derived 4'-phosphopantetheine (4'PP) cofactor to the side chain hydroxyl group of the absolutely conserved serine residue in the core T-domain structural motif LGG(HD)**S**(LI) (Marahiel et al., 1997). This converts the T-domain from a non-functional apo-protein to a functional holo-protein able to accept aminoacyl-AMP substrates from the upstream A-domain, which is bound as an aminoacyl-adduct via a thioester bond to the 4'PP cofactor (Lambalot et al., 1996; Reuter et al., 1999). Solved structures for bacterial T-domains have demonstrated these carrier domains consist of a bundle of four α-helices (Koglin et al., 2006; Lohman et al., 2014), with the hydrophobic interactions between the third α-helix and the C-terminal domain of the 4'PPTase required for cofactor attachment (Tufar et al., 2014). The flexible 4'PP cofactor then enables transfer of the aminoacyl adduct to upstream or

downstream condensation domains to facilitate peptide bond formation (Lohman et al., 2014).

1.2.3 Condensation domains

NRPS condensation (C) domains catalyse peptide bond formation between the donor aminoacyl or peptidyl group attached to the T-domain 4'PP cofactor of the preceding module and the acceptor aminoacyl group attached to the T-domain 4'PP cofactor from the same module (Hur et al., 2012). This results in release of donor moiety from the preceding module, with the resulting peptide chain retained on the acceptor T-domain 4'PP cofactor (Finking and Marahiel, 2004). C-domains, which are approx. 450 aa in length, are pseudo-dimers with the active site located within a solvent channel at the junction between N- and C-terminal subunits (Keating et al., 2002). The solvent channel allows access to the active site by the 4'PP cofactor of both the donor and acceptor T-domains (Keating et al., 2002). Studies of bacterial proteins have shown that C-domains usually exhibit strong stereoselectivity towards both donor and acceptor substrates, and generally also exhibit side-chain specificity towards the acceptor substrate (Lautru and Challis, 2004).

1.2.4 *N*-Methylation domains

NRPS *N*-Methylation (M) domains are usually nested between the conserved A8 and A9 motifs of a host AMPylation domain (Ansari et al., 2008), and are one of several types of NRPS accessorising domains that modify the amino acid substrate of their respective modules (Hur et al., 2012). These M-domains transfer methyl groups from *S*-adenosyl-methionine (SAM) donors onto the α-amino nitrogen atom of the 4'PP-bound amino acid substrate prior to peptide bond formation (Hur et al., 2012). *N*-methylation has been shown to stabilise the active conformation of non-ribosomal peptides, contribute to bioactivity and protect against proteolytic cleavage (Marahiel et al., 1997; Velkov et al., 2011). Studies have shown that while the product synthesis efficiency may be lowered, blocking substrate methylation often does not prevent peptide bond formation, thereby resulting in synthesis of a non-methylated or partially-methylated products (Glinski et al., 2001; Velkov et al., 2011).

1.2.5 Termination domains

Termination domains are located at the C-terminus of the final module in most NRPS megasynthases, and are responsible for release of the final non-ribosomal peptide (NRP) product (Du and Lou, 2010). In bacteria this function is predominately fulfilled by type I thioesterase (TE) domains, which release NRP products via hydrolysis or macrocyclisation reactions (Du and Lou, 2010). However, TE-domain-mediated NRP release is rare in fungi, with only a few examples described to date (Balibar et al., 2007; Eisfeld, 2009). Product termination for many fungal NRPSs is catalysed by C-terminal condensation domains, which catalyse release of a cyclic peptide product (Eisfeld, 2009). Another common fungal NRPS release mechanism is provided by reductase (R) domains, which catalyse the NAD(P)H-dependent reduction of thioester-bound NRP products to form a terminal alcohol or aldehyde group (Du and Lou, 2010). R-domains exhibit a Rossmann-fold structural motif with a NAH(P)H binding site and Ser-Tyr-Lys catalytic triad, similar to short-chain reductase (SDR) superfamily proteins (Du and Lou, 2010). An R-domain variant has also been identified in which product release is catalysed by a non-redox mechanism similar to a Dieckmann condensation reaction, resulting in product release via an intramolecular cyclisation reaction that results in formation of a 5- or 6-membered ring structure (Liu and Walsh, 2009).

1.2.6 Type II thioesterases

Type II thioesterases (TEII) are α/β-hydrolase proteins required to maintain efficient product synthesis by NRPS and PKS megasynthases (Kotowska and Pawlik, 2014). These stand-alone proteins are described as “type II” to differentiate them from type I TEs, which are product release domains integrated within PKS and NRPS enzymes. The promiscuity of 4’PPTase enzymes is such that acyl-CoA donors are often used as substrate for post-translational modification of NRPS T-domains (Finking and Marahiel, 2004). This results in T-domains with non-reactive acyl-4’PP cofactors that are unable to bind amino acid substrates during product synthesis. TEIIs catalyse the hydrolytic release of such thioester-linked acyl moieties, regenerating the T-domain 4’PP cofactor, and are also able to regenerate 4’PP cofactors primed with aberrant aminoacyl substrates (Koglin et al., 2008; Schwarzer

et al., 2002; Yeh et al., 2004). Inefficient TEII-mediated release of a D-Phe–Pro–Leu tripeptide from a T-domain has also been demonstrated *in vitro* (Schwarzer et al., 2002), though other studies have shown that peptidyl-4'PP groups tend to be resistant to TEII cleavage (Koglin et al., 2008; Yeh et al., 2004). This preference for short-chain acyl substrates is proposed to be due to the shallow substrate-binding cleft of TEII proteins relative to TE domains, which contain deeper clefts to accommodate the NRPS/PKS peptidyl product prior to release (Koglin et al., 2008). Nevertheless, product release roles have been demonstrated for TEIIs in some PKS/NRPS systems (Kotowska and Pawlik, 2014). Bacterial TEIIs are usually found as part of NRPS/PKS-based SM clusters (Kotowska and Pawlik, 2014). However, there are very few examples of fungal NRPS/PKS clusters with associated TEII enzymes (Zabala et al., 2014). This is probably because fungal genomes usually encode many more NRPS/PKS proteins than smaller bacterial genomes (Bushley and Turgeon, 2010; Wang et al., 2014). Fungal TEIIs may therefore be required to service many different NRPS/PKS proteins, while a bacterial TEII generally only services the NRPS/PKS encoded by the cluster within which is it located (Kotowska and Pawlik, 2014).

1.2.7 Adenylate-forming reductases

The biosynthetic pathway for L-lysine in fungi uniquely involves an adenylate-forming reductase (AFR) protein for conversion of α-aminoadipic acid into α-aminoadipate-6-semialdeyde (Bhattacharjee, 1985; Casqueiro et al., 1998). These AFR proteins contain a NRPS-like A-T-R domain architecture, yet are not involved in peptide synthesis (Casqueiro et al., 1998). Recent studies have identified AFR proteins in fungi not involved in the lysine biosynthetic pathway, including L-serine, L-tyrosine and 5-methyl orsellinic acid AFRs (Forseth et al., 2013; Kalb et al., 2014). Interestingly, the functionally redundant *Aspergillus flavus* LnaA and LnbA AFR proteins are proposed to catalyse formation of a dimeric imine from their reduced L-tyrosine products, rather than releasing these as monomeric aldehyde products as observed for other AFRs (Forseth et al., 2013). This may indicate an unexpected diversity in potential products for AFR proteins.

1.2.8 NRPS genes in *Epichloë* spp.

Johnson et al. (2007b) identified that *Epichloë* spp. contain at least 12 different NRPS-encoding genes using a degenerate PCR approach, and a survey of the genomes now published for many *Epichloë* spp. would likely identify more (Schardl et al., 2013a). These NRPS genes show variable distribution across *Epichloë*, with some conserved across most or all *Epichloë* spp., while the distribution of others is quite discontinuous (Johnson et al., 2007b). These NRPS genes are predicted to have been inherited from a common ancestor of all *Epichloë* spp., with gene loss resulting in the diversity observed today (Johnson et al., 2007b). The most well-studied of these *Epichloë* genes are *lpsA* and *lpsB*, which encode NRPSs involved in ergot alkaloid biosynthesis (Fleetwood et al., 2007; Haarmann et al., 2005; Young et al., 2015), and *perA*. The *perA* gene encodes the peramine synthetase protein PerA, which is a two-module NRPS (Tanaka et al., 2005). This NRPS catalyses peptide bond formation between two amino acid substrates, proposed to be 1-pyrroline-5-carboxylate (P5C) for the first module, and arginine for the second module (Tanaka et al., 2005). PerA also methylates the α-amino nitrogen atom of arginine, and contains a C-terminal non-redox type R-domain that catalyses an intramolecular cyclisation between the arginine carboxylate and P5C amino group to form the distinctive pyrrolopyrazine double ring structure and releasing the dipeptide product from PerA (Tanaka et al., 2005). Spontaneous rearrangement and air oxidation steps are then predicted to result in the final product, peramine (Tanaka et al., 2005). This predicted biosynthetic pathway is detailed in Figure 1.4.

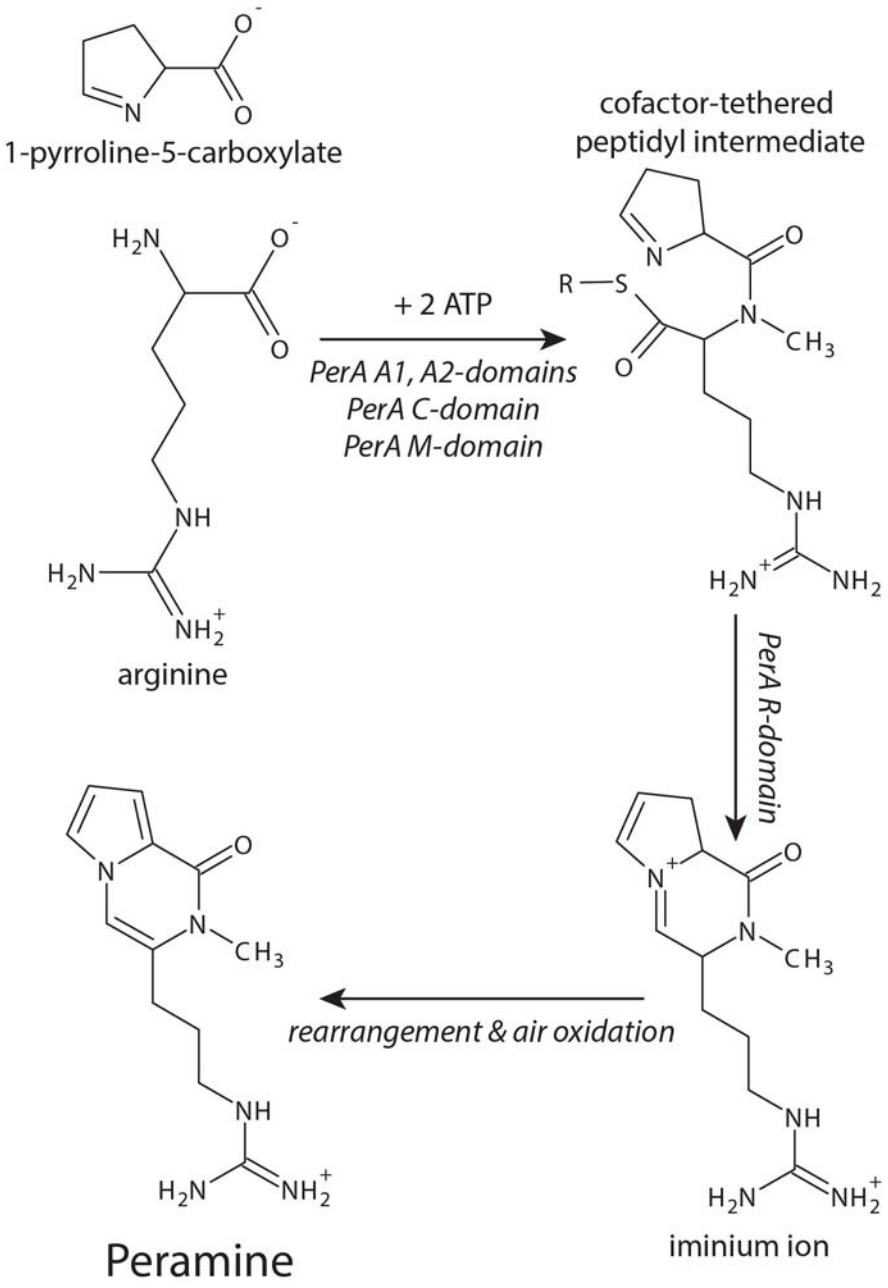


Figure 1.4 Peramine biosynthetic pathway as proposed by Tanaka et al., (2005). Catalytic domains of the two-module NRPS PerA catalyse all biosynthetic steps except the final rearrangement and air oxidation, which are proposed to be spontaneous.

1.3 *E. festucae* Fl1–*Lolium perenne* as a model symbiosis

The mutualistic benefits of asexual (type III) *Epichloë* strains means that they are the predominant endophytes used in association with pastoral grasses (Johnson et al., 2013a; Young et al., 2013). *Lolium perenne* (perennial ryegrass) is a forage grass used in temperate areas with moderate to high rainfall, and is the most common pastoral grass in New Zealand (Johnson et al., 2013a; Young et al., 2013). The asexual non-hybrid *E. festucae* var. *lolii* is the most common naturally occurring endophyte in *L. perenne*, although the sexual *E. festucae* progenitor of this variant (Leuchtmann et al., 2014) can also be readily inoculated into *L. perenne* seedlings to form a stable, systemic and asymptomatic synthetic association (Scott et al., 2007). *E. festucae* var. *lolii* is very slow growing in culture and particularly intransigent towards genetic modification. This makes it unsuitable for genetic analysis of the interaction between *Epichloë* spp. and their grass hosts. Instead, the synthetic association between *L. perenne* and *E. festucae* Fl1 [originally isolated from *Festuca trachyphylla* (hard fescue; Siegel et al., 1990)]has been established as a model for this symbiotic interaction over the past two decades (Scott et al., 2007). Fl1 has a number of desirable traits for a model *Epichloë* endophyte; it is relatively fast growing, easy to transform, produces a wide variety of bio-protective secondary metabolites and only rarely forms stromata on host plants, simplifying containment.

In recent years both targeted gene deletions and forward genetic screens have significantly increased our knowledge of the molecular mechanisms underpinning this *E. festucae* Fl1–*L. perenne* symbiosis, with implications for *Epichloë*/Poöideae interactions in general. These advances include elucidation of the SM genes and pathways responsible for production of the ergot alkaloids, indole-diterpenes and peramine (Fleetwood et al., 2007; Tanaka et al., 2005; Young et al., 2005; Young et al., 2006). Additionally, a variety of symbiosis-defective *E. festucae* Fl1 gene deletion strains have been generated in *noxA*, encoding a NADPH oxidase (Tanaka et al., 2006); *sakA*, encoding a stress-activated mitogen-activated protein kinase (Eaton et al., 2008); *proA*, encoding a C6-Zn transcription factor (Tanaka et al., 2013); *mpkA* or *mkkA*, which encode a mitogen activated protein (MAP) kinase and kinase kinase of the cell-wall integrity signalling pathway, respectively (Becker et al.,

2015); and others (Charlton et al., 2012; Johnson et al., 2013b; Takemoto et al., 2011; Takemoto et al., 2006; Tanaka et al., 2008). These mutants typically exhibit phenotypes such as host stunting, hyper-proliferation of endophytic hyphae and colonisation of vascular bundles, all of which are atypical for *E. festucae* Fl1. Many of these mutants were found to be deficient in cell-cell fusion, leading to the hypothesis that an interconnected hyphal network is required to maintain symbiotic growth *in planta* (Becker et al., 2015; Charlton et al., 2012; Kayano et al., 2013).

This rich resource of symbiosis-defective gene deletion mutants has enabled the application of RNAseq to compare the transcriptomes from *L. perenne* plants infected with *E. festucae* Fl1 wt, Δ*sakA*, Δ*noxA* or Δ*proA* (Eaton et al., 2010; Eaton et al., 2015), and more recently Δ*hepA* (Chujo et al., unpublished data) to identify genes that are differentially expressed across all mutant associations. Transcription of *perA* and the *IDT*/*EAS* cluster genes, which are all expressed *in planta* (Fleetwood et al., 2007; Tanaka et al., 2005; Young et al., 2005; Young et al., 2006), was commonly identified as downregulated in these symbiotic mutants, particularly Δ*sakA*, along with a number of uncharacterised SM genes (Eaton et al., 2010; Eaton et al., 2015). This indicates that downregulated genes in these symbiosis-defective associations are candidates for encoding products with symbiosis-related functions, although results from targeted gene deletion mutants based on this data have yet to be published. Eaton et al. (2010) identified six differentially regulated gene clusters in Δ*sakA*, three of which were downregulated. These downregulated clusters included the *IDT* cluster, a cluster including the symbiotically regulated gene *gigA* (Johnson et al., 2003; Johnson et al., 2007a; Johnson et al., 2015) and a cluster predicted to encode genes involved in nitrogen metabolism (Eaton et al., 2010). One gene from this putative nitrogen metabolism gene cluster encoding a putative NmrA-like protein was subsequently found to be significantly downregulated across *ΔsakA*, Δ*noxA*, Δ*proA* (Eaton et al., 2015) and Δ*hepA* (Chujo et al., unpublished results), one of fewer than 40 genes to exhibit such regulation. NmrA is an *Aspergillus nidulans* protein that mediates post-translational repression of the alternative nitrogen metabolism pathway-activating transcription factor AreA (Stammers et al., 2001).

1.4 Aims

There were two main aims that the research presented in this thesis focused on. The first aim was based on observations indicating that the presence of a *perA* gene was not a particularly reliable indicator for the capacity of an isolate to produce peramine. The aim was therefore to identify and characterise the mutations responsible for abolishing peramine production in *perA* genes, with a particular focus on the widespread *perA-ΔR* allele. This aim was divided into the following objectives:

1. **Perform a survey of *perA* alleles across *Epichloë*.** This objective was addressed by characterisation of diverse *perA* alleles though PCR and gene sequencing analyses.
2. **Design a simple screen for determining the likely peramine chemotype of an *Epichloë* isolate.** This objective was addressed by utilising the results of objective 1 to determine a screening protocol utilising PCR and gene sequence data.
3. **Establish *Penicillium paxilli* as a heterologous expression host for *perA* alleles.** This objective was initially addressed through sequencing the genome of *P. paxilli* to confirm its suitability for expression of NRPS proteins. The *E. festucae* Fl1 *perA* allele was then heterologously expressed in *P. paxilli* to confirm peramine production.
4. **Investigate the A1-domain substrate specificity of PerA-ΔR proteins.** This objective was addressed through replacement of the *E. festucae* Fl1 *perA* region that encodes the A1-domain substrate-binding residues with the equivalent region from several *perA-ΔR* alleles. These constructs were then heterologously expressed in *P. paxilli* and assayed for peramine production.
5. **Reconstitute peramine production for the *E. festucae* E2368 *perA-ΔR* allele to provide insight into the functionality of PerA-ΔR proteins.** This objective was addressed through swapping domain-encoding sequences between this *perA-ΔR* allele and a *perA* allele known to encode a protein capable of producing peramine. The resulting constructs were then heterologously expressed in *P. paxilli* and assayed for peramine production.

6. **Investigate product release mechanisms for PerA-ΔR proteins.** This objective was achieved through heterologous expression of truncated *perA* alleles with *perA* or *perA-ΔR*-type T2-domains. These constructs were heterologously expressed in *P. paxilli* with or without an *E. festucae* TEII-encoding gene to investigate the requirements for product release from PerA-ΔR proteins.
7. **Investigate if PerA-ΔR proteins produce a novel product.** This objective was achieved through heterologous expression of *perA-ΔR* in *P. paxilli* with or without an *E. festucae* TEII-encoding gene to determine if a novel product synthesised by PerA-ΔR proteins could be identified.

The second aim of this study was to characterise the gene of interest *irlA*, which was the "*nmrA*-like" gene identified as part of a differentially regulated nitrogen metabolism gene cluster in a transcriptome comparison of wt vs. Δ*sakA E. festucae* in *L. perenne* hosts (Eaton et al., 2010). This aim was addressed by the following objectives:

8. **Determine if *irlA* is important for the *E. festucae–L. perenne* symbiosis.** This objective was addressed by targeted gene deletion of *irlA* in *E. festucae* Fl1, followed by analysis of gene deletion mutants *in planta*.
9. **Identify the function of the IrlA protein.** This objective was addressed though bioinformatic analysis of *irlA* and the genomic locus within which it is located. This identified *irlA* as part of a putative five-gene SM cluster that differed from the cluster predicted by Eaton et al. (2010).
10. **Investigate the function of the five-gene cluster within which *irlA* is located.** The function of this cluster was examined through using bioinformatics tools, RT-PCR analysis of gene expression in different plant tissues and gene deletion analyses.

2 Materials and Methods

2.1 Preface

Unless stated otherwise, all solutions and media were made up with Milli-Q purified water and sterilised by autoclaving. All work with live cells or material destined for inclusion with live cell cultures was performed under sterile conditions. Materials and methods for Berry et al. (2015a) and Berry et al. (2015b) are located within those publications.

2.2 Molecular and biological material

Table 2.1 Details of strains used in this study

Strain	Relevant characteristics	Source or reference
E. festucae		
PN2278	wtild type(Fl1)	(Young et al., 2005)
PN3040	Fl1/Δ*irlA*::P*trpC–hph*; Hyg^R (Δ*irlA*-62)	This study
PN3041	Fl1/Δ*irlA*::P*trpC–hph*; Hyg^R (Δ*irlA*-64)	This study
PN3042	Fl1/Δ*irlA*::P*trpC–hph*; Hyg^R (Δ*irlA*-133)	This study
PN3043	Fl1/Δ*irlA*::P*trpC–hph*; Hyg^R (Δ*irlA*-172)	This study
PN3110	Fl1/Δ*irlA*::P*trpC–hph*; Hyg^R (Δ*irlA*-438)	This study
PN3111	Fl1/Δ*irlA*::P*trpC–hph*; Hyg^R (Δ*irlA*-487)	This study
PN3112	Fl1/Δ*irlA*::P*trpC–hph*; Hyg^R (Δ*irlA*-573)	This study
08-1	Δ*irlA*/*irlA*; PN3040/pDB08, pII99; Hyg^R, Gen^R	This study
08-2	Δ*irlA*/*irlA*; PN3040/pDB08, pII99; Hyg^R, Gen^R	This study
08-3	Δ*irlA*/*irlA*; PN3040/pDB08, pII99; Hyg^R, Gen^R	This study
09-1	Δ*irlA*/*irlA-eGFP*; PN3040/pDB09, pII99; Hyg^R, Gen^R	This study
09-2	Δ*irlA*/*irlA-eGFP*; PN3040/pDB09, pII99; Hyg^R, Gen^R	This study
09-3	Δ*irlA*/*irlA-eGFP*; PN3040/pDB09, pII99; Hyg^R, Gen^R	This study
PN3109	Fl1/Δ*pdtA*::P*trpC–hph*; Hyg^R (Δ*pdtA*-168)	This study
PN3080	Fl1/Δ*afrA*::P*trpC–hph*; Hyg^R (Δ*afrA*-11)	This study
PN3081	Fl1/Δ*mfsB*::P*trpC–hph*; Hyg^R (Δ*mfsB*-153)	This study
PN3082	Fl1/Δ*fxbA*::P*trpC–hph*; Hyg^R (Δ*fxbA*-15)	This study
P. paxilli		
PN2013	wild type (ATCC 26601)	(Itoh et al., 1994)
PN2253	PN2013/Δ*paxT–paxD*; Hyg^R; paxilline negative	(Young et al., 2001)
ΔPAX05-47	Δ*paxT–paxD*/P*paxM–perA*; PN2253/pDB05/pII99; Hyg^R, Gen^R	This study
ΔPAX05-49	Δ*paxT–paxD*/P*paxM–perA*; PN2253/pDB05/pII99; Hyg^R, Gen^R	This study
ΔPAX05-55	Δ*paxT–paxD*/P*paxM–perA*; PN2253/pDB05/pII99; Hyg^R, Gen^R	This study
05-31	PN2013/P*paxM–perA*; pDB05/pAN7-1; Hyg^R	This study
05-33	PN2013/P*paxM–perA*; pDB05/pAN7-1; Hyg^R	This study
05-43	PN2013/P*paxM–perA*; pDB05/pAN7-1; Hyg^R	This study
15-2	PN2013/P*paxM–perA-ΔR*; pDB15/pSF15.15; Hyg^R	This study
15-4	PN2013/P*paxM–perA-ΔR*; pDB15/pSF15.15; Hyg^R	This study
15-5	PN2013/P*paxM–perA-ΔR*; pDB15/pSF15.15; Hyg^R	This study
15/34-12	PN2013/P*paxM–perA-ΔR*, P*tefA–tesB*; pDB15/pDB34; Hyg^R	This study
15/34-36	PN2013/P*paxM–perA-ΔR*, P*tefA–tesB*; pDB15/pDB34; Hyg^R	This study

15/34-57	PN2013/P*paxM–perA-ΔR*, P*tefA–tesB*; pDB15/pDB34; Hyg^R	This study
15/35-4	PN2013/P*paxM–perA-ΔR*, P*tesB–tesB*; pDB15/pDB34; Hyg^R	This study
15/35-8	PN2013/P*paxM–perA-ΔR*, P*tesB–tesB*; pDB15/pDB34; Hyg^R	This study
15/35-21	PN2013/P*paxM–perA-ΔR*, P*tesB–tesB*; pDB15/pDB34; Hyg^R	This study
17-5	PN2013/P*paxM–perA-17*; pDB05/pSF15.15; Hyg^R	This study
17-16	PN2013/P*paxM–perA-17*; pDB17/pSF15.15; Hyg^R	This study
17-18	PN2013/P*paxM–perA-17*; pDB17/pSF15.15; Hyg^R	This study
18-7	PN2013/P*paxM–perA-18*; pDB18/pSF15.15; Hyg^R	This study
18-9	PN2013/P*paxM–perA-18*; pDB18/pSF15.15; Hyg^R	This study
18-15	PN2013/P*paxM–perA-18*; pDB18/pSF15.15; Hyg^R	This study
18/34-16	PN2013/P*paxM–perA-18*, P*tefA–tesB*; pDB18/pDB34; Hyg^R	This study
18/34-20	PN2013/P*paxM–perA-18*, P*tefA–tesB*; pDB18/pDB34; Hyg^R	This study
18/34-36	PN2013/P*paxM–perA-18*, P*tefA–tesB*; pDB18/pDB34; Hyg^R	This study
18/35-6	PN2013/P*paxM–perA-18*, P*tesB–tesB*; pDB18/pDB35; Hyg^R	This study
18/35-11	PN2013/P*paxM–perA-18*, P*tesB–tesB*; pDB18/pDB35; Hyg^R	This study
18/35-16	PN2013/P*paxM–perA-18*, P*tesB–tesB*; pDB18/pDB35; Hyg^R	This study
19-1	PN2013/P*paxM–perA-19*; pDB19/pSF15.15; Hyg^R	This study
19-11	PN2013/P*paxM–perA-19*; pDB19/pSF15.15; Hyg^R	This study
19-19	PN2013/P*paxM–perA-19*; pDB19/pSF15.15; Hyg^R	This study
19/34-2	PN2013/P*paxM–perA-19*, P*tefA–tesB*; pDB19/pDB34; Hyg^R	This study
19/34-10	PN2013/P*paxM–perA-19*, P*tefA–tesB*; pDB19/pDB34; Hyg^R	This study
19/34-18	PN2013/P*paxM–perA-19*, P*tefA–tesB*; pDB19/pDB34; Hyg^R	This study
19/35-3	PN2013/P*paxM–perA-19*, P*tesB–tesB*; pDB19/pDB35; Hyg^R	This study
19/35-17	PN2013/P*paxM–perA-19*, P*tesB–tesB*; pDB19/pDB35; Hyg^R	This study
19/35-20	PN2013/P*paxM–perA-19*, P*tesB–tesB*; pDB19/pDB35; Hyg^R	This study
20-2	PN2013/P*paxM–perA-20*; pDB20/pSF15.15; Hyg^R	This study
20-5	PN2013/P*paxM–perA-20*; pDB20/pSF15.15; Hyg^R	This study
20-14	PN2013/P*paxM–perA-20*; pDB20/pSF15.15; Hyg^R	This study
29-1	PN2013/P*paxM–perA-29*; pDB29/pSF15.15; Hyg^R	This study
29-14	PN2013/P*paxM–perA-29*; pDB29/pSF15.15; Hyg^R	This study
29-15	PN2013/P*paxM–perA-29*; pDB29/pSF15.15; Hyg^R	This study
30-2	PN2013/P*paxM–perA-30*; pDB30/pSF15.15; Hyg^R	This study
30-6	PN2013/P*paxM–perA-30*; pDB30/pSF15.15; Hyg^R	This study
30-28	PN2013/P*paxM–perA-30*; pDB30/pSF15.15; Hyg^R	This study
31-4	PN2013/P*paxM–perA-31*; pDB31/pSF15.15; Hyg^R	This study
31-7	PN2013/P*paxM–perA-31*; pDB31/pSF15.15; Hyg^R	This study
31-9	PN2013/P*paxM–perA-31*; pDB31/pSF15.15; Hyg^R	This study

32-2	PN2013/P*paxM–perA-32*; pDB32/pSF15.15; Hyg^R	This study
32-5	PN2013/P*paxM–perA-32*; pDB32/pSF15.15; Hyg^R	This study
32-21	PN2013/P*paxM–perA-32*; pDB32/pSF15.15; Hyg^R	This study
43-2	PN2013/P*paxM–perA-43*; pDB43/pSF15.15; Hyg^R	This study
43-15	PN2013/P*paxM–perA-43*; pDB43/pSF15.15; Hyg^R	This study
43-20	PN2013/P*paxM–perA-43*; pDB43/pSF15.15; Hyg^R	This study
34-4	PN2013/P*tefA–tesB*; pDB34; Hyg^R	This study
34-5	PN2013/P*tefA–tesB*; pDB34; Hyg^R	This study
34-6	PN2013/P*tefA–tesB*; pDB34; Hyg^R	This study
35-1	PN2013/P*tesB–tesB*; pDB35; Hyg^R	This study
35-2	PN2013/P*tesB–tesB*; pDB35; Hyg^R	This study
35-3	PN2013/P*tesB–tesB*; pDB35; Hyg^R	This study
426-15	PN2013; pRS426/pSF15.15; Hyg^R	This study
426-16	PN2013; pRS426/pSF15.15; Hyg^R	This study
426-17	PN2013; pRS426/pSF15.15; Hyg^R	This study
S. cerevisiae		
FY834	S288C/*MAT*α *his3*Δ*200 ura3-52 leu2*Δ*1 lys2*Δ*202 trp1*Δ*63*	(Winston et al., 1995)
E. coli		
DH5α	F^- Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rk^-, mk^+) *pho*A*sup*E44 λ^- *thi*$^-$1 *gyr*A96 *rel*A1	ThermoFisher Scientific
DH5α-E	F^- Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rk^-, mk^+) *pho*A*sup*E44 λ^- *thi*$^-$1 *gyr*A96 *rel*A1	ThermoFisher Scientific
K12 (*dam*$^-$ /*dcm*$^-$)	*ara-14 leuB6 fhuA31 lacY1 tsx78 glnV44 galK2 galT22 mcrA dcm-6 hisG4 rfbD1 R(zgb210::Tn10)* Tet^S *endA1 rspL136* (Str^R) *dam13::Tn9* (Cam^R) *xylA-5 mtl-1 thi-1 mcrB1 hsdR2*	New England Biolabs
DB02	DH5α-E/pDB02	This study
DB05	DH5α-E/pDB05	This study
DB08	DH5α/pDB08	This study
DB09	DH5α/pDB09	This study
DB15	DH5α/pDB15	This study
DB17	DH5α/pDB17	This study
DB18	DH5α/pDB18	This study
DB19	DH5α/pDB19	This study
DB20	DH5α/pDB20	This study
DB24	DH5α/pDB24	This study
DB25	DH5α/pDB25	This study
DB26	DH5α/pDB26	This study

DB27	DH5α/pDB27	This study
DB29	DH5α/pDB29	This study
DB30	DH5α/pDB30	This study
DB31	DH5α/pDB31	This study
DB32	DH5α/pDB32	This study
DB34	DH5α/pDB34	This study
DB35	DH5α/pDB35	This study
DB36	DH5α/pDB36	This study
DB43	DH5α/pDB43	This study

Table 2.2 Plasmids used in this study[1]

Plasmid	Relevant characteristics	Details[1]	Source or Reference
pRS426	Amp^R	9.3.1	
pSF15.15	P*trpC–hph*; Amp^R, Hyg^R	9.3.2	
pAN7-1	P*gpdA–hph*–T*trpC*; Amp^R, Hyg^R	9.3.3	
pII99	P*trpC–nptII*–T*trpC*; Amp^R, Hyg^R	9.3.4	(Namiki et al., 2001)
pDB02	*irlA* deletion construct; Amp^R, Hyg^R	9.3.5	This study
pDB05	P*paxM–perA*–T*paxM*; Amp^R	9.3.6	This study
pDB08	P*irlA–irlA*–T*irlA*; Amp^R	9.3.7	This study
pDB09	P*irlA–irlA*:*eGFP*–T*irlA*; Amp^R	9.3.8	This study
pDB15	P*paxM–perA-ΔR*–T*paxM*; Amp^R	9.3.9	This study
pDB17	P*paxM–perA-17*–T*paxM*; Amp^R	9.3.10	This study
pDB18	P*paxM–perA-18*–T*paxM*; Amp^R	9.3.11	This study
pDB19	P*paxM–perA-19*–T*paxM*; Amp^R	9.3.12	This study
pDB20	P*paxM–perA-20*–T*paxM*; Amp^R	9.3.13	This study
pDB24	*pdtA* deletion construct; Amp^R, Hyg^R	9.3.14	This study
pDB25	*afrA* deletion construct; Amp^R, Hyg^R	9.3.15	This study
pDB26	*mfsB* deletion construct; Amp^R, Hyg^R	9.3.16	This study
pDB27	*fxbA* deletion construct; Amp^R, Hyg^R	9.3.17	This study
pDB29	P*paxM–perA-29*–T*paxM*; Amp^R	9.3.18	This study
pDB30	P*paxM–perA-30*–T*paxM*; Amp^R	9.3.19	This study
pDB31	P*paxM–perA-31*–T*paxM*; Amp^R	9.3.20	This study
pDB32	P*paxM–perA-32*–T*paxM*; Amp^R	9.3.21	This study
pDB34	P*tefA–tesB*(*Efes*)–T*tesB*(*Ppax*)::P*trpC–hph*–T*trpC*; Amp^R, Hyg^R	9.3.22	This study
pDB35	P*tesB*(*Ppax*)–*tesB*(*Efes*)–T*tesB*(*Ppax*)::P*trpC–hph*–T*trpC*; Amp^R, Hyg^R	9.3.23	This study
pDB36	*irlA* deletion construct; Amp^R, Hyg^R	9.3.24	This study
pDB43	P*paxM–perA-43*–T*paxM*; Amp^R	9.3.25	This study

[1] Plasmid maps located in appendix at address indicated by the "Details" column.

Table 2.3 Primers used in this study

Primer	Designed for	Primer Sequence
24SCR_3	amplify across *pdtA* deletion construct left border	CGGTTGGCATGACAAACCTG
24SCR_4	amplify across *pdtA* deletion construct right border	TGCTCTTCTCAGCAGTGACG
24SCR_5	amplify within *pdtA*	ACATCGCACAAGGGAGCTAC
24SCR_6	amplify within *pdtA*	GCATGCTTGGAAGTTCTGGC
25SCR_1	amplify within *afrA*	AACTGGACCAGTATGCCACG
25SCR_2	amplify within *afrA*	CACGCATATGTGGCATTCCG
25SCR_3	amplify across *afrA* deletion construct left border	TGGTATCGGCTCTCCAAACG
25SCR_4	amplify across *afrA* deletion construct right border	GTTGACGCTCTGATGTTGCC
26SCR_3	amplify across *mfsB* deletion construct left border	AAGGACAGCTTGACAGTCGG
26SCR_4	amplify across *mfsB* deletion construct right border	TCTTCGCAGATCCATCGTCG
26SCR_5	amplify within *mfsB*	TCAAATGCGGCACAAAGAGC
26SCR_6	amplify within *mfsB*	TTGCAGTTGTTCGCCACAAC
27SCR_3	amplify across *fxbA* deletion construct left border	GATGCTTACCACGCCAATGG
27SCR_4	amplify across *fxbA* deletion construct right border	TTTCGAGGAGAAGCTCGTGC
27SCR_5	amplify within *fxbA*	TGGGCATCTGGTGATTGACC
27SCR_6	amplify within *fxbA*	AAACTTCCGCTCCAGTACGG
afrA_3	amplify within *afrA* cDNA (RT-qPCR)	GTTTGTCTACGTGTCCAACCG
afrA_4	amplify within *afrA* cDNA (RT-qPCR)	ACCGAGTCTACGTGTACCAGA
afrA_4.1	amplify within *afrA* cDNA (RT-qPCR)	ACCGAGTCCACGTGTACCAGA
Dgl_1	amplify within cDNA from *D. glomerata* 5.8S rRNA (RT-qPCR)	TATCTCGGCTCTCGCATCGATG
Dgl_2	amplify within cDNA from *D. glomerata* 5.8S rRNA (RT-qPCR)	AAAGACTCGATGGTTCGCGG
Dgl_5	amplify within cDNA from *D. glomerata* cytosolic heat shock protein 90.2 gene(RT-qPCR)	AAGAGGAAGGAGGAGCTCAAGG
Dgl_6	amplify within cDNA from *D. glomerata* cytosolic heat shock protein 90.2 gene(RT-qPCR)	AATCACCTTCTCCACCCTGTCG
EF320_1	amplify within *EF320* cDNA (RT-qPCR)	AGCGTTGTGGTTTCCGCATC
EF320_2	amplify within *EF320* cDNA (RT-qPCR)	TCGGCAAGGTGGCATTTTCC
esdC_1	amplify within *esdC* cDNA (RT-qPCR)	ACGGAGATGACAGCTCTGAC
esdC_2	amplify within *esdC* cDNA (RT-qPCR)	ACCACTTCCCAGCACATCAG
fxbA_3	amplify within *fxbA* cDNA (RT-qPCR)	TGAAACTCACGGCCATTATCC

fxbA_3.1	amplify within *fxbA* cDNA (RT-qPCR)	TGAAACTCACGGCCACTATCC
fxbA_4	amplify within *fxbA* cDNA (RT-qPCR)	CTTCTGTGCCGTCTCTGAATC
fxbA_4.1	amplify within *fxbA* cDNA (RT-qPCR)	CTTCTGTGCCGTCTTTGAATC
fxbA_5	amplify within *fxbA* cDNA (RT-qPCR)	AAGAGAGCTCAACCTGTGGC
fxbA_6	amplify within *fxbA* cDNA (RT-qPCR)	GACATCTGCGATTGCTGAGC
GRA4	amplify within *tesB* cDNA (RT-PCR)	TCTTCACAAACACGCGCAAG
GRA5	amplify within *tesB* cDNA (RT-PCR)	ATGGTCTGAGTCGGATGACG
hph_iF	amplify from P*trpC* towards 5' direction	CCTTCAATATCAGTTCCAAGCT
hph_iR	amplify from *hph* towards 3' direction	CGTCCGAGGGCAAAGGAATAG
hyg_F	amplification of P*trpC–hph* cassette	AGCTTGGAACTGATATTGAAGG
hyg_R	amplification of P*trpC–hph* cassette	CTATTCCTTTGCCCTCGGACG
irlA_3	amplify within *irlA* cDNA (RT-qPCR)	CAAGTGGCGTACTGAGCATCT
irlA_4	amplify within *irlA* cDNA (RT-qPCR)	ATGGTTCCCAGCACAAACACA
irlA_locus_F	amplification of *irlA* locus	ACCTTGTTCGACATTGCACCAC
irlA_locus_R	amplification of *irlA* locus	CAGGGAAATTTCTGTCGCCACC
irlA_scr_1	amplify across *irlA* deletion construct left border	CATGGACGGTATAACCTGGG
irlA_scr_2	amplify across *irlA* deletion construct left border	ATGTCCTCGTTCCTGTCTGC
irlA_scr_3	amplify across P*trpC–hph* cassette in *irlA* deletion construct	ATCAGCTCGGTGAGATTTCG
irlA_scr_4	amplify across P*trpC–hph* cassette in *irlA* deletion construct	GGCTGTGTCGTACACATTGG
irlA_scr_5	amplify across *irlA* deletion construct right border	TGTCTGGCGTACACAAATCG
irlA_scr_6	amplify across *irlA* deletion construct right border	GCTCGACAAGAAGACGTTGG
KG1	screen for P*tesB* integration	ATCCACTGCTCCAAGATCCC
KG11	amplify 3' end of *perA* cDNA (RT-PCR)	GTTTCTAGTGCGCGATGC
KG12	amplify 3' end of *perA* cDNA (RT-PCR)	CTAGCCTCCAGATCTTGTG
KG13	amplify 3' end of *perA-ΔR* cDNA (RT-PCR)	TGTTCCGATTTCCCACTG
KG14	amplify 3' end of *perA-ΔR* cDNA (RT-PCR)	TCAACATCGGGCAAGGTC
KG2	screen for P*tefA* and P*tesB* integration	ACTTACCGTGCCTAGAACCC
KG8	screen for P*tefA* integration	TTAGCAACAGGCCAGGCTAG
mfs_5.1	amplify within *mfsB* cDNA (RT-qPCR)	TGAAACTGCCGACCTGTTCG
mfsB_3	amplify within *mfsB* cDNA (RT-qPCR)	ATTGGATGCCTCGTACCTCTG
mfsB_4	amplify within *mfsB* cDNA (RT-qPCR)	TCGTCTCCTTCTGCGTCATAC
mfsB_4.1	amplify within *mfsB* cDNA (RT-	TCGTCTTGTTCTGCGTCATAC

	qPCR)	
mfsB_5	amplify within *mfsB* cDNA (RT-qPCR)	TAAGACTGCCGACCTGTTCG
mfsB_6	amplify within *mfsB* cDNA (RT-qPCR)	GTTACGGACAAGATTGGCGC
P+GFP_R	amplification of P*tefA–eGFP* cassette	TTACTTGTACAGCTCGTCCATGC
p9f	amplify within A2 domain-encoding region of *perA*	GCAAACGCCGTCTCTGCTCA
p9r	amplify within A2 domain-encoding region of *perA*	GGATCCCCTTAACAACCACT
pDB02_1	synthesis of pDB02	GTAACGCCAGGGTTTTCCCAGTCACGACGAA TTCCAACGAGTCGGTGAGAGACA
pDB02_2	synthesis of pDB02	ATGCTCCTTCAATATCAGTTCCAAGCTAAGG TCACCGTTTAGTTGGCT
pDB02_3	synthesis of pDB02	CCAGCACTCGTCCGAGGGCAAAGGAATAGGC TGCAGCTTCCTATCATCG
pDB02_4	synthesis of pDB02	GCGGATAACAATTTCACACAGGAAACAGCGA ATTCTAATGCTGCTGCAGAGGAAC
pDB05_1	synthesis of pDB05	GTAACGCCAGGGTTTTCCCAGTCACGACAAG CTTGTTGTTGGCATGGGAGTAGG
pDB05_10	synthesis of pDB05	GCGGATAACAATTTCACACAGGAAACAGCAA GCTTCGAATTGAGAAGCTGATTTGC
pDB05_2	synthesis of pDB05	TCAACCTCATCAAAAGGCTCCGCGTCCATGG TTTCTGAATCTTAAAGATACATG
pDB05_3	synthesis of pDB05	ATGGACGCGGAGCCTTTTG
pDB05_4	synthesis of pDB05	TACGCTACAACCCTCATCG
pDB05_5	synthesis of pDB05	AATAGGCTCCTCGTTCTCC
pDB05_6	synthesis of pDB05	ATCGTGAGCAAGTGATTGG
pDB05_7	synthesis of pDB05	TCTGACTGACCAGGAGCTG
pDB05_8	synthesis of pDB05	CTAGCCTCCAGATCTTGTG
pDB05_9	synthesis of pDB05	TGCGTGCTTTCACAAGATCTGGAGGCTAGAC CATTGGAGCAATTTTTGG
pDB08_1	synthesis of pDB08	GTAACGCCAGGGTTTTCCCAGTCACGACAAG CTTTCCATCTGTCCAGATCGTGG
pDB08_2	synthesis of pDB08	GCGGATAACAATTTCACACAGGAAACAGCAA GCTTGGCAATGTGCACGAAGAACC
pDB09_1	synthesis of pDB09	CCGGTGAACAGCTCCTCGCCCTTGCTCACGG CGCCGGCGCCGGCGCCTGTAGCATTCACGTT CCCC
pDB09_3	synthesis of pDB09	CTCTCGGCATGGACGAGCTGTACAAGTAATT CATGTTCACAGCCGCTGC
pDB15_1	synthesis of pDB15	TCGGCGGAGAACGAGGAGC
pDB15_2	synthesis of pDB15	AAGCCACAGCTCCTGGTCAG
pDB15_3	synthesis of pDB15	GGAAAACCAAAAATTGCTCCAATGGTTCAAC ATCGGGCAAGGTCTG
pDB15_4	synthesis of pDB15	ACCATTGGAGCAATTTTTGG
pDB15_5	synthesis of pDB15	GGTTTCTGAATCTTAAAGATACATG
pDB17_1	synthesis of pDB17	CAATGGCACCATCGGAAACG
pDB17_2	synthesis of pDB17	CGTCGACCTTTCCCCACCATCGGGTAGCATC GTCCACAGACAGAGGCTCGCCGGAAAACACC ATCGTCTTTAGAAAAGG
pDB17_3	synthesis of pDB17	CCCGATGGTGGGGAAAGGTCGACGTCGTCAA CGGATATGGGCCTGCAGAGTGCATCATCAAC ACTGTCAACAGCC
pDB17_4	synthesis of pDB17	AAGTCGGTTGCCTTTTTGGG

pDB18_1	synthesis of pDB18	CCATCCACTCCGAACTTGC
pDB18_2	synthesis of pDB18	GGAAAACCAAAAATTGCTCCAATGGTTCAAC ATCGGGCAAGGTCCGACAG
pDB19_1	synthesis of pDB19	TCAGGACCTTTGACCAGATTTGC
pDB19_2	synthesis of pDB19	ATGCAGCAAATCTGGTCAAAGG
pDB20_1	synthesis of pDB20	TGAGAGTCGCAAACACGTTATG
pDB20_2	synthesis of pDB20	TCCGTTCATAACGTGTTTGC
pDB20_3	synthesis of pDB20	AGCAGCTTCAGGACTGATAGG
pDB20_4	synthesis of pDB20	AACAGCCGACCTATCAGTCCTG
pDB24_1	synthesis of pDB24	ACGCCAGGGTTTTCCCAGTCACGACGTTAAC CTCTGGCTTTGTGACTTTGG
pDB24_2	synthesis of pDB24	CACTCGTCCGAGGGCAAAGGAATAGCACCTT GGTTTGCAATGTAG
pDB24_3	synthesis of pDB24	GCTCCTTCAATATCAGTTCCAAGCTGAATTG CGCTCGTTTATCATC
pDB24_4	synthesis of pDB24	ATAACAATTTCACACAGGAAACAGCCCCGGG CAGTGCGTCGAAGAGCTGC
pDB25_1	synthesis of pDB25	ACGCCAGGGTTTTCCCAGTCACGACGTTAAC TGGCAGGATCTTTGTAGAGG
pDB25_2	synthesis of pDB25	GCTCCTTCAATATCAGTTCCAAGCTAGCTTC AATGGTGCTGACAG
pDB25_3	synthesis of pDB25	CACTCGTCCGAGGGCAAAGGAATAGTCGGCT TGAGTGTGATGTAG
pDB25_4	synthesis of pDB25	ATAACAATTTCACACAGGAAACAGCCCCGGG AATCGAAGTCCAGGTGAACG
pDB26_1	synthesis of pDB26	ACGCCAGGGTTTTCCCAGTCACGACGTTAAC CAGGCACAATCACCAAAGTC
pDB26_2	synthesis of pDB26	CACTCGTCCGAGGGCAAAGGAATAGGGTGAA GTAGGTCCGAGATC
pDB26_3	synthesis of pDB26	GCTCCTTCAATATCAGTTCCAAGCTAGACGT TGATAACAATGCTG
pDB26_4	synthesis of pDB26	ATAACAATTTCACACAGGAAACAGCCCCGGG TACAAACGGAATCTGCTTGG
pDB27_1	synthesis of pDB27	ACGCCAGGGTTTTCCCAGTCACGACGTTAAC GACCAGCCTTGCGATGACC
pDB27_2	synthesis of pDB27	GCTCCTTCAATATCAGTTCCAAGCTCAAGAT GAGCACACCCAAGG
pDB27_3	synthesis of pDB27	CACTCGTCCGAGGGCAAAGGAATAGATTCGT CCTCTGCCCTAATC
pDB27_4	synthesis of pDB27	ATAACAATTTCACACAGGAAACAGCCCCGGG AGAGGACGGAGAGAATCAAG
pDB29_1	synthesis of pDB29	ACCACAGAGCCAATCACCTGCTCACGATGAG ACACAATTAAAGC
pDB29_2	synthesis of pDB29	TCTACATCCTGGCGAATCG
pDB30_2	synthesis of pDB30	TCCAGGCTTTCATAAAGTAC
pDB30_3	synthesis of pDB30	GCTTCCTGTGTGATGGCCTG
pDB30_5	synthesis of pDB30	TGGACGTACTTTATGAAAGCC
pDB30_6	synthesis of pDB30	ATCTCCAGGCCATCACACAG
pDB31_1	synthesis of pDB31	GAGGGCGGAAGAGTAGTTGG
pDB31_2	synthesis of pDB31	CGCACACCAACTACTCTTCC
pDB31_3	synthesis of pDB31	AAGCAGTTCCCCAACACAGC
pDB31_4	synthesis of pDB31	AACCATCAAGTACTCGTTCC
pDB34_1	synthesis of pDB34	ACGCCAGGGTTTTCCCAGTCACGACGGATCC GGTAGCAAACGGTGGTCAAAGG
pDB34_2	synthesis of pDB34	GCGTGGACCGTTCCTCGCGAGCCATGAATTC

		GAGCTCGGTACCC
pDB34_3	synthesis of pDB34	ATGGCTCGCGAGGAACGGTC
pDB34_4	synthesis of pDB34	ATAGTATTAGGTAGGGTATAATGGGTTACAA CTTTGTCTTTTCCTTTTCAG
pDB34_5	synthesis of pDB34	CCCATTATACCCTACCTAATAC
pDB34_6	synthesis of pDB34	GCTCCTTCAATATCAGTTCCAAGCTTCACGG AGATGCAAAGTTGG
pDB35_1	synthesis of pDB35	ACGCCAGGGTTTTCCCAGTCACGACGGATCC ACTGCTCCAAGATCCCTAAG
pDB35_2	synthesis of pDB35	GCGTGGACCGTTCCTCGCGAGCCATGATGGT TTCGGAGTATCTCAAG
pDB36_1	synthesis of pDB36	GCGGATAACAATTTCACACAGGAAACAGCTG ATCATAATGCTGCTGCAGAGGAAC
pdtA_3	amplify within *pdtA* cDNA (RT-qPCR)	AACTTCGGGTTGCTCAATACC
pdtA_4	amplify within *pdtA* cDNA (RT-qPCR)	AGAGGTCCGAGTTTTGGCAC
pdtA_4.1	amplify within *pdtA* cDNA (RT-qPCR)	AAAGGTCCGAGTTTTGGCAC
pdtA_5	amplify within *pdtA* cDNA (RT-qPCR)	TACCCTCCCGTTTGCAACTC
pdtA_6	amplify within *pdtA* cDNA (RT-qPCR)	CGTTTCTAGCCGTTTCTGCG
perA1_4	amplify within *perA* cDNA (RT-PCR)	TCGGAAAGGTCGGCTGTAC
perA1_R	amplify within *perA* cDNA (RT-PCR)	TTGCTTCATCCCAGTCAGC
PPperA_SCR_1	screening for P*paxM* integration	TGTTTCTGGTACGTCGACCG
PPperA_SCR_2	screening for P*paxM* integration	CTAGAGGCGCATAGGACACG
PPperA_SCR_3	screening for T*paxM* integration	AGCGCTGTCAGATATGCTCC
PPperA_SCR_4	screening for T*paxM* integration	CAGAGTTGTCCGGGAAGTCC
pRS426_F	amplify vector backbone for Gibson assembly	GCTGTTTCCTGTGTGAAATTG
pRS426_R	amplify vector backbone for Gibson assembly	GTCGTGACTGGGAAAACCC
pRS426_SCR_F	amplify within pRS426	TAGAGCTTGACGGGGAAAGC
pRS426_SCR_R	amplify within pRS426	ATAGTCCTGTCGGGTTTCGC
TC399	amplify within *EF2* cDNA (RT-qPCR)	AAAAAGCAACCGAATGCAAG
TC399.1	amplify within *EF2* cDNA (RT-qPCR)	TAAAAGCAACCGAATGCAAG
TC400	amplify within *EF2* cDNA (RT-qPCR)	CGAGACGACATAACTACATGTATCAAA
TC403.1	amplify within *rps10* cDNA (RT-qPCR)	GCATGGGCAAACACAACCTA
TC404.1	amplify within *rps10* cDNA (RT-qPCR)	CAGATTTGCCCTCCAATTTT
TC407	amplify within *rps22* cDNA (RT-qPCR)	TAGCTGGCGTTATGGAAAGG
TC408	amplify within *rps22* cDNA (RT-qPCR)	CGATTGTGCGACTACTACCTCA
TC408.1	amplify within *rps22* cDNA (RT-qPCR)	CGATTGTGCGACTGCCTCAACG

tefA_1	amplify within *tefA* cDNA (RT-qPCR)	CGACATTGCCCTCTGGAAGT
tefA_2	amplify within *tefA* cDNA (RT-qPCR)	CAGCAGCGATAATCAGGATAG
tub2_F	amplify within *P. paxilli tub2* cDNA (RT-qPCR)	ACGGCCTTGACGGTGATGGA
tub2_R	amplify within *P. paxilli tub2* cDNA (RT-qPCR)	GAACTCCTCACGGATCTTGG

2.3 Growth media

2.3.1 Czapek Dox-yeast extract + trace elements (CDYE + TE) media

3.34% Czapek Dox (Oxoid; w/v), 0.5% yeast extract (w/v) and 5% trace element solution (2.3.6; v/v)

2.3.2 Lysogeny broth (LB) medium (Miller, 1972)

1% tryptone (w/v), 0.5% yeast extract (w/v) and 86 mM NaCl; 1.5% agar (w/v) added for solid media (referred to as LBA). Adjusted to pH 7 with NaOH.

2.3.3 Potato dextrose (PD) medium

2.4% potato dextrose (w/v); 1.5% agar (w/v) added for solid media (referred to as PDA).

2.3.4 SOC medium

2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4 \cdot 7H_2O$ and 20 mM glucose.

2.3.5 YPD medium

1% (w/v) yeast extract, 2% mycological peptone, pH 5.8. Add water to 950 mL, autoclave, then add 50 mL of sterile 40% (w/v) glucose. 2% (w/v) agar added before autoclaving for solid media.

2.3.6 Trace element solution

1.7 mM $FeSO_4 \cdot 7H_2O$, 1.73 mM $ZnSO_4 \cdot 7H_2O$, 0.59 mM $MnSO_4 \cdot H_2O$, 0.2 mM $CuSO_4 \cdot 5H_2O$ and 0.17 mM $CoCl_2 \cdot 6H_2O$ made up in 0.6 M HCl.

2.4 Growth and storage of microbiological cultures

2.4.1 *Escherichia coli* (*E. coli*)

2.4.1.1 Liquid culture

E. coli liquid cultures were routinely grown in sterile 20 × 150 mm (diameter × height) test tubes containing 4 mL liquid LB medium (2.3.2), supplemented with antibiotics when appropriate (2.4.1.3), for approx. 18 h at 37°C with shaking of 200 RPM. Inoculations were initiated by transfer of a solid culture colony (2.4.1.2) or fluid from a thawed cryopreservation culture using a sterilised wire loop (2.4.1.4), or by pipetting 2 µL from a previous liquid culture.

2.4.1.2 Solid culture

E. coli solid cultures were routinely grown in sterile 90 × 15 mm (diameter × height) circular petri dishes containing approx. 25 mL solid LB medium (2.3.2), supplemented with antibiotics where appropriate (2.4.1.3), for approx. 18 h at 37°C. Plated cultures were stored at 4°C for up to a month as required.

2.4.1.3 Antibiotic concentrations

Filter-sterilised antibiotics were added to liquid or molten media under sterile conditions post-autoclaving to a final concentration of 100 µg mL^{-1} for ampicillin or 50 µg mL^{-1} for kanamycin.

2.4.1.4 Cryopreservation

A lawn of bacteria was grown overnight on solid media (2.4.1.2) containing suitable antibiotic selection (2.4.1.3). A 5 mL volume of 50% glycerol (v/v) was then dispensed onto the surface of this media, and a glass spreader was used to resuspend the bacteria. This suspension was dispensed in 1 mL aliquots into 1.8 mL cryotubes and stored at -80°C.

2.4.2 *Epichloë festucae* (*E. festucae*)

2.4.2.1 Liquid culture

E. festucae liquid cultures were routinely grown in 125 mL conical flasks containing 50 mL PD liquid medium (2.3.3), supplemented with antibiotics when appropriate (2.4.2.3), for 4-5 days at 22°C with shaking of 200 RPM. Inoculum was approx. 1 cm^2 of freshly grown mycelia finely diced using a sterile scalpel and scraped off the surface of a solid culture colony (2.4.2.2).

2.4.2.2 Solid culture

E. festucae solid cultures were routinely grown in sterile 90 × 15 mm (diameter × height) circular petri dishes containing approx. 25 mL solid PDA medium (2.3.3), supplemented with antibiotics when appropriate (2.4.2.3), for 7-10 days at 22°C. Plated cultures were stored at 4°C for up to six months as required.

2.4.2.3 Antibiotic concentrations

Antibiotics were added to liquid or molten media under sterile conditions post-autoclaving to a final concentration of 150 µg mL^{-1} for hygromycin or 200 µg mL^{-1} for geneticin.

2.4.2.4 Single spore isolation

Colonies of *E. festucae* were grown on solid media (2.4.2.2) for 10 d, then 2 mL of liquid PD medium (2.3.3) used to flood each plate, and the colonies were scrubbed with a sterilised glass spreader. The resulting suspension was decanted and filtered through 1 mL pipette tips filled with glass wool into 1.5 mL microtubes. The filtrate was then spread in 400 µL aliquots onto fresh solid media plates and incubated at 22°C until the single spore-derived colonies were visible.

2.4.3 *Penicillium paxilli*

2.4.3.1 Liquid culture

P. paxilli liquid cultures were routinely grown in sterile 125 mL conical flasks containing 50 mL CDYE + TE liquid medium (2.3.1), supplemented with antibiotics

when appropriate (2.4.3.3). Unless stated otherwise, cultures were grown for 5 d at 22°C with shaking of 200 RPM. An inoculum of 5×10^6 spores was used for each flask.

2.4.3.2 Solid culture

P. paxilli solid cultures were routinely grown in sterile 90 × 15 mm (diameter × height) circular petri dishes containing approx. 25 mL solid CDYE + TE medium (2.3.3), supplemented with antibiotics when appropriate (2.4.3.3), for 5-7 days at 22°C. Plated cultures were stored at 4°C for up to six months as required.

2.4.3.3 Antibiotic concentrations

Antibiotics were added to liquid or molten media under sterile conditions post-autoclaving to a final concentration of 100 µg mL^{-1} for hygromycin or 150 µg mL^{-1} for geneticin.

2.4.3.4 Preparation of spore suspensions

Square blocks of conidiating mycelia (approx. 2 cm^2) were cut from a solid culture of *P. paxilli* (2.4.3.2) and placed in a sterile 10 mL Bijou bottle containing 4 mL of 0.01% Triton X-100. This bottle was then shaken vigorously for 5 s to suspend the spores, and the colloidal suspension was then transferred to sterile microtubes. The spore concentration was determined using a haemocytometer, and the suspension was stored at 4°C for up to a month.

2.4.3.5 Single spore isolation technique

A sterilised loop was dipped into a spore suspension (2.4.3.4; spores resuspended by gentle agitation) then streaked onto a petri dish containing solid media (2.3.1) containing an appropriate antibiotic for selection (2.4.3.3) to achieve single spore isolation. Plates were incubated at 22°C until colonies appeared, with a single colony selected for subsequent experiments.

2.4.4 Harvesting fungal mycelia from liquid cultures

After growing for the desired length of time, fungal liquid cultures (2.4.2.1, 2.4.3.1) were harvested by pouring the entire culture through a sterile coarse filter (e.g. nappy liner) contained within a glass funnel. If desired, filtered media samples were taken from the flow-through. The filtered mycelia were washed twice with 200

mL water, with excess water squeezed through the filter. The isolated mycelia were scraped off the filter using a stainless steel spatula for use in subsequent applications. This procedure was performed in a sterile environment using autoclaved materials and reagents when live cells were required, e.g. protoplast generation (2.9.1).

2.5 DNA isolation techniques

2.5.1 Plasmid isolation from *E. coli* liquid cultures

High quality plasmid DNA was routinely isolated from *E. coli* liquid cultures (2.4.1.1) using the Wizard® *Plus* SV Miniprep system (Promega), as per the manufacturer's instructions.

2.5.2 Plasmid screening for *E. coli* transformation-derived colonies

Colonies from *E. coli* transformations (2.9.3, 2.9.4) were screened after overnight growth by picking individual colonies using sterile 20 µL pipette tips and resuspending the bacteria in 0.2 mL PCR tubes containing 6 µL of LB media (2.3.2), with antibiotic selection suitable for maintaining the desired plasmid (2.4.1.3). A 3 µL aliquot of this suspension was transferred to a second 0.2 mL PCR tube containing 8 µL of lysis solution (0.1 M NaCl, 10 mM Tris-HCl, 1 mM EDTA and 0.5% Triton X-100, adjusted to pH 8), mixed by pipetting and then heated to 99°C for 30 s in a thermocycler. The remaining 3 µL of bacterial suspension was stored at 4°C. After cooling the lysis solution to room temperature, a 2 µL solution containing 2-10 units of the desired restriction enzyme and 1 µL of the corresponding 10× reaction buffer was added to this lysate, mixed by pipetting, then incubated at 37°C (or at the temperature appropriate for the chosen restriction enzyme) for 30 min. The digested plasmid was then visualised (2.6.8), and transformants producing the desired digest pattern were inoculated for subsequent experiments from the remaining bacterial suspension stored earlier at 4°C.

2.5.3 Crude plasmid extraction from *Saccharomyces cerevisiae*

Yeast cells cultured on solid media were resuspended in 1 mL YPD medium (2.3.5) by scrubbing with a sterile glass spreader. The yeast suspension was

transferred to a 2 mL microtube and the scrubbing process was repeated once more, pooling both cell suspensions. The cells were pelleted by centrifugation at 17,000 × g for 15 s and the supernatant was discarded. The pellet was lysed with 100 µL chloroform, 100 µL phenol, 100 µL lysis buffer [2% (v/v) Triton X-100, 1% (w/v) SDS, 100 mM NaCl, 10 mM Tris (from 1 M stock solution at pH 8) and 1 mM EDTA] and approx. 100 µL of 400-650 µm glass beads by vigorous vortexing for 2 min. After centrifugation at 17,000 × g for 10 min a 100 µL aliquot of the supernatant was removed to a new 1.5 mL microtube, mixed with 10 µL 3 M sodium acetate and 250 µL 95% (v/v) ethanol, and incubated for 10 min at room temperature. After centrifugation at 17,000 × g for 10 min the supernatant was decanted and the pellet washed with 1 mL 70% (v/v) ethanol. After air-drying for approx. 10 min the pellet was resuspended in 50 µL water and 1 µL was used for electroporation of *E. coli* (2.9.4).

2.5.4 Isolation of genomic DNA (gDNA) from fungal liquid cultures

This method for small scale isolation of gDNA from fungal liquid cultures was adapted from Byrd et al. (1990). Fungal mycelia were isolated (2.4.4) from liquid culture with normal growing times of 3-4 d for *P. paxilli* (2.4.3.1) or 4-5 d for *E. festucae* (2.4.2.1), then lyophilised for at least 24 h. Approx. 20 mg of lyophilised mycelia was ground to a fine powder under liquid nitrogen using a mortar and pestle, then resuspended in 800 µL of extraction buffer (150 mM EDTA, 50 mM Tris-HCl and 1% (w/v) SDS, adjusted to pH 8). This lysate was transferred into a 2 mL microtube to which proteinase K was added to a final concentration of 2 mg mL^{-1}, followed by incubation at 37°C for 20 min. The sample was centrifuged for 10 min at 17,000 × g, following which the supernatant was transferred to a new 2 mL microtube. Chloroform and phenol were added in equal volumes of 400 µL to this tube and mixed thoroughly by vortexing. Following centrifugation for 10 min at 17,000 × g the aqueous upper phase was carefully removed by pipetting and transferred to a new tube, and the phenol/chloroform extraction step was repeated twice more using the aqueous phase each time. A final extraction step was done using 800 µL chloroform, then 800 µL of isopropanol was added to the separated aqueous phase and mixed thoroughly by inversion. The DNA was then pelleted by

centrifugation at 17,000 × g for 10 min, the supernatant was discarded, and the pellet was washed with 1 mL of 70% (v/v) ethanol followed by centrifugation for 5 min at 17,000 × g. The supernatant was discarded, and the pellet was air-dried at room temperature for 15 min then resuspended in 100 µL Milli-Q purified H_2O and stored at 4°C.

2.5.5 Isolation of crude gDNA from fungal colonies

E. festucae fungal mycelia were grown on solid media for approx. 7 d (2.4.2.2). Approx. 1 cm^2 of mycelia were chopped finely using a scalpel then scraped off the petri dish and placed in a microtube containing 150 µL of lysis buffer (100 mM EDTA, 100 mM Tris-HCl and 1% (w/v) SDS, adjusted to pH 8). The tubes were incubated at 70°C for 30 min, then 150 µL of 5 M potassium acetate was added to each tube, which were then inverted 8 times and placed on ice for 10 min. Tubes were centrifuged for at least 20 min at 17,000 × g, then a 200 µL sample of the supernatant was transferred to a fresh microtube, with care taken not to disturb the pelleted material. 140 µL of isopropanol was added to each of these tubes, which were inverted 8 times to mix and centrifuged for 15 min at 17,000 × g. The supernatant was discarded and the pellet was washed with 180 µL 70% (v/v) ethanol and centrifuged for 5 min at 17,000 × g. The supernatant was removed by pipetting and the pellet was allowed to air dry for 10 min. The pellet was resuspended in 50 µL of Milli-Q purified H_2O and heated at 70°C for 10 min, centrifuged at 17,000 × g for 10 min and finally transferred to a new tube or 96 well plate before being used in downstream applications or stored at -20°C.

2.5.6 Isolation of crude gDNA from fungal spores

Spore suspensions (2.4.3.4) were resuspeded by gentle agitation, then a 200 µL sample was removed to a 2 mL screw-cap microtube containing 100 µL of 425-600 µm glass beads. These tubes were shaken in a FastPrep® FP120 Cell Disrupter System (Thermo Savant) for 20 s at 4 min/s. Cellular debris and glass beads were then pelleted by centrifugation at 4°C/14,000 × g for 10 min, and a 100 µL supernatant sample was removed to a new microtube and used as template for PCR.

2.5.7 High-throughput DNA isolation from infected plant material

Total plant DNA (including endophyte) was extracted using the MagAttract 96 DNA plant core kit (Qiagen, Inc., Valencia, CA) as per the manufacturer's instructions.

2.6 DNA manipulation techniques

2.6.1 Digestion with restriction enzymes

Restriction enzyme digestion of DNA was performed using enzymes from Roche and New England Biolabs. Digestions were generally performed as per the manufacturer's instructions. However, 1.1 µg samples of gDNA for Southern analysis were digested in a total volume of 200 µL (approx. 20:1 dilution of the gDNA) for 4 h with 20 U of enzyme with the appropriate buffer and at the appropriate temperature. An additional 10 U of enzyme was then added before overnight incubation. Double digests were only performed when a suitable buffer for both enzymes was available, with DNA otherwise purified between sequential digestions.

2.6.2 Ethanol precipitation of nucleic acids

DNA was precipitated by adding 0.1 volume of 3 M sodium acetate and 2.5 volumes of absolute ethanol to a 1.5 mL microtube containing the DNA in solution and mixing by inversion. Samples were centrifuged for 10 min at 17,000 × g and the supernatant was discarded. 1 mL 70% (v/v) ethanol was then added to rinse the pellet, and the sample was centrifuged for 5 min at 17,000 × g. The supernatant was discarded, and the pellet was air-dried in a fume hood for 10 min at room temperature prior to resuspension in the desired volume of water.

2.6.3 Standard PCR amplifications

Standard PCR reactions were performed using Taq DNA polymerase (Roche) or OneTaq® DNA polymerase (New England Biolabs, Inc.), with reaction conditions as per the manufacturer's instructions. 2% DMSO (v/v) was added to the PCR reaction mix for difficult to amplify templates.

2.6.4 High-fidelity PCR amplifications

Phusion® DNA polymerase (ThermoFisher Scientific) or Q5® DNA polymerase (New England Biolabs, Inc.) were used for PCR when high-fidelity products were

required for downstream applications, with reaction conditions as per the manufacturer's instructions.

2.6.5 DNA purification from PCR reactions

DNA was purified from PCR reactions using the Wizard® SV Gel and PCR Clean-UP System (Promega), as per the manufacturer's instructions. Gel purification (2.6.9) was used when gel electrophoresis (2.6.8) indicated the presence of non-specific PCR products.

2.6.6 Quantitative PCR (qPCR) using Roche LC480

Reaction mixes for qPCR using the LightCycler® 480 II (Roche) were prepared under clean conditions using filter tips from 2.5 µL water, 0.25 µL of each primer from 10 µM stock solutions, 2 µL of appropriately diluted DNA or cDNA template and 5 µL of 2× SsoFast™ EvaGreen® Supermix (Bio-Rad) for a final volume of 10 µL. Reactions were then run for 40 cycles (95°C for 5 s, 60°C for 20 s). Melting curves were captured after the 40 amplification cycles by ramping the temperature from 65°C to 95°C at a rate of 0.11°C s^{-1}. Data was analysed using the LightCycler® 480 software, version 1.5 (Roche).

2.6.7 Quantitative PCR (qPCR) using ABI 7900HT

Reaction mixes for qPCR using the 7900HT real-time PCR system (Applied Biosystems) were prepared under clean conditions using filter tips from 2.5 µL water, 0.25 µL of each primer from 10 µM stock solutions, 2 µL of appropriately diluted DNA or cDNA template and 5 µL of 2× Power Syber Green master mix (ThermoFisher Scientific) for a final volume of 10 µL. Reactions were then run for 40 cycles (95°C for 15 s, 60°C for 60 s). Melting curves were captured after 40 amplification cycles by ramping the temperature from 60°C to 95°C at a 2% ramp rate. Data was analysed using the 7900 SDS software, version 2.4 (Applied Biosystems).

2.6.8 Gel electrophoresis

Gels were prepared from molten agarose (0.7% w/v for separation of large DNA fragments and up to 2% w/v for separation of smaller fragments) prepared with

1x TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM Na_2EDTA). Solidified gels were immersed in 1x TBE buffer and loaded with DNA samples mixed 5:1 with 6 × loading dye [20% sucrose (w/v), 5 mM EDTA, 1% (w/v) SDS and 0.2% (w/v) bromophenol blue) and run using electrophoresis equipment. The gel was then stained for 10–15 min with ethidium bromide (1 µg mL^{-1} ethidium bromide) and destained for 10 min in tap water (Milli-Q filtered water used for gel purifications). Gels were visualized and photographed using a UV Transilluminator Gel Documentation System (Bio-Rad).

2.6.9 Gel purification

Purification of DNA fragments from agarose gels was done using the Wizard® SV Gel and PCR Clean-UP System (Promega), as per the manufacturer's instructions.

2.6.10 DNA ligation

Ligation reactions were performed using 20 ng of linearised vector DNA that had been dephosphorylated by incubation with rAPid alkaline phosphatase (Roche), as per the manufacturer's instructions. A 3:1 molar excess of purified insert:vector DNA in a final volume 30 µL was incubated with 2 U of T4 ligase (Roche) at 4°C overnight. This ligation reaction mix was then used to transform chemically competent or electrocompetent *E. coli* cells (2.9.3; 2.9.4).

2.6.11 Yeast recombinational cloning

The vector for yeast recombinational cloning was prepared via restriction enzyme digestion (2.6.1) of pRS426 using *Eco*RI and *Xho*I followed by gel purification (2.6.9) of a 5.5-kb fragment. Insert fragments were prepared via high-fidelity PCR (2.6.4) and gel purification (2.6.9). These fragments were designed to contain 40-bp homologous overlaps at the desired recombination sites. 5 mL of YPD liquid media (2.3.5) in a capped test tube was inoculated from a single *S. cerevisiae* strain FY834 colony using a sterilised wire loop. This culture was grown overnight at 30°C/200 RPM, then 1 mL of this culture was pipetted into 50 mL YPD liquid media in a 125 mL conical flask and grown at 30°C/200 RPM for 4 h, or until the OD was between 0.4 and 0.5. The resulting culture was centrifuged in a 50 mL universal tube at 1200 × g for 5 min. The supernatant was discarded and the cells were washed with 25 mL water, centrifuged at 1200 × g, resuspended in 1 mL 100 mM lithium acetate and

transferred to a 1.5 mL microtube. The cells were pelleted again by centrifugation at 17,000 × g for 15 s and resuspended in 400 µL of 100 mM lithium acetate. A 50 µL aliquot of this suspension was transferred to a new 1.5 mL microtube and centrifuged at 17,000 × g for 15 s, following which the supernatant was discarded and the cell pellet was kept on ice. The carrier DNA (2 mg mL^{-1} salmon or herring sperm DNA) was boiled for 5 min, placed on ice, then 10 µL of the chilled carrier DNA was mixed with 240 µL 50% (w/v) PEG 4000 and 36 µL 1 M lithium acetate and kept on ice. This mix was added to the cell pellet along with 34 µL of DNA solution containing approx. 300 ng of each fragment to be assembled. Cells were resuspended by gentle vortexing, then incubated at 30°C for 30 min. Cells were then pelleted at 17,000 × g for 15 s and washed with 1 mL water three times, and resuspended in 50 µL of water. Equal aliquots of this suspension were spread across two SD ura^- plates [18% (w/v) sorbitol, 0.67% (w/v) yeast nitrogen base without amino acids (Sigma-Aldrich), 0.077% (w/v) uracil dropout supplement (Clontech), 2% (w/v) agar; made up to 950 µL, adjusted pH to 5.8, autoclaved, then 50mL of sterile 40% (w/v) glucose was added] These plates were incubated at 30°C for 3 d or until colonies appeared, following which single colonies were streaked onto solid media (2.3.5) and grown at 30°C for 3 d for plasmid extraction (2.5.3).

2.6.12 Gibson assembly (Gibson et al., 2009)

Gibson assembly master mix (2× concentration) was prepared on ice by combining 100 µL of 5× isothermal buffer [25% (w/v) PEG-8000, 500 mM Tris-HCl (from 1 M pH 7.5 stock solution), 50 mM $MgCl_2$, 50 mM 1,4-dithiothreitol, 5 mM NAD and 1 mM dNTPs] with 2 µL T5 exonuclease (10 U $µL^{-1}$; New England Biolabs, Inc.), 6.25 µL Phusion DNA polymerase (2 U $µL^{-1}$; ThermoFisher Scientific), 50 µL *Taq* ligase (40 U $µL^{-1}$; New England Biolabs, Inc.) and 91.75 µL of water to a final volume of 250 µL. Aliquots (25 × 10 µL) were placed into PCR tubes and stored at -20°C until required. DNA fragments for use in Gibson assembly reactions were generated by high-fidelity PCR (2.6.4) and designed so that fragments to be joined contained homologous overlaps that were at least 15-bp long with a T_m >50°C, with no additional sequence beyond those overlaps. DNA fragments for assembly were then purified (2.6.5; 2.6.9) and combined to a final volume of 10 µL containing 50-100 ng

of vector DNA and each insert fragment with 3:1 relative molar concentration to the vector. This mixture was then combined on ice with a 10 µL aliquot of Gibson assembly 2× master mix, incubated in a thermocycler at 50°C for 1 h, and then used to transform chemically competent *E. coli* cells (2.9.3).

2.6.13 DNA sequencing

DNA sequencing of plasmids and PCR products was performed by the Massey Genome Service using a 3730 DNA Analyser (Applied Biosystems) with BigDye™ Terminator Version 3.1 chemistry (Applied Biosystems).

2.7 RNA isolation and manipulation

Workspaces and utensils intended for use with RNA extractions were carefully cleaned, then treated with RNase AWAY® (Sigma-Aldrich) to remove residual RNase contamination. All pipetting was done using RNase AWAY®-cleaned pipettes with filter tips, and only guranteed RNase-free consumables were used.

2.7.1 RNA isolation using TRIzol®

Fungal mycelial samples for RNA extraction were harvested (2.4.4) then immediately frozen in liquid nitrogen. Infected grass samples for RNA extractions were harvested by removing dead plant material, sectioning into 10 mm lengths then immediately frozen in liquid nitrogen. Frozen mycelial/plant samples were ground to a fine powder under liquid nitrogen using a mortar and pestle, then 1 mL TRIzol® (Thermo Fisher Scientific) was added and the sample mixed until a frozen paste formed. This paste was allowed to thaw and was transferred to a 2 mL microtube and centrifuged at 4°C/10,000 × g for 10 min. The supernatant was transferred to a new 1.5 mL microtube, to which 200 µL of chloroform was added and mixed by vortexing, the tube was centrifuged at 4°C/10,000 × g for 15 min. The aqueous upper phase was transferred to a new 1.5 mL microtube along with 500 µL isopropanol, and the sample inverted 6 times then centrifuged at 4°C/10,000 × g for 10 min. The supernatant was discarded and the pellet was washed with 75% ethanol and air-dried in a negative-flow cabinet. DEPC-treated water (100 µL) was added to the pellet and resuspended by pipetting and incubation at 70°C for 10 min. The RNA

sample was then treated with RQ1 RNase-free DNase (Promega) as per the manufacturer's instructions and stored at -80°C until required.

2.7.2 RNA isolation using the Qiagen RNeasy Plant Mini Kit

RNA isolations using the RNeasy Plant Mini Kit (Qiagen) were performed as per the manufacturer's instructions.

2.7.3 Reverse transcription

Synthesis of cDNA from 0.4 – 1 µg of RNA was performed with the QuantiTect® Reverse Transcription Kit (Qiagen), as per the manufacturer's instructions.

2.8 Preparation and probing of Southern blots

2.8.1 Preparation of Southern blots

High quality gDNA (1 µg; 2.5.1) from each sample was digested using the desired restriction enzymes (2.6.1). These digests were mixed 5:1 with 6× loading dye and loaded into each well of a large format 0.7% (w/v) agarose gel. Digested samples were concentrated by ethanol precipitation (2.6.2) if volume reduction was required for gel loading. Gels were run overnight at 30 V, stained with ethidium bromide, destained in water and imaged under UV next to a reference measure. Gels were depurinated in 0.25 M HCl for 15 min, denaturated (0.5 M NaOH, 0.5 M NaCl) for 40 min and then neutralized (0.5 M Tris, pH 7.4, 2 M NaCl) for 40 min before being washed in 2× SSC (0.3 M NaCl, 30 mM sodium citrate) for 2 min and placed on a blotting apparatus for overnight DNA transfer to a positively-charged nylon membrane (Roche). The blotting apparatus consisted of 2 "wicks" of 320 x 156 mm Whatman 3MM paper soaked in 20× SSC (3 M NaCl, 0.3 M sodium citrate), the ends of which were submerged in wells of 20× SSC. The gel was then placed on top, with a protective plastic cling-film layer with a square hole cut in the centre slightly smaller than the size of the gel place between wick and gel to prevent a capillary action short-circuit. The positively charged nylon transfer membrane soaked in 2× SSC cut 4 mm larger than the gel was placed on top, followed by 2 sheets of Whatman 3MM paper of the same size as the gel soaked in 2× SSC, and two more identical dry

sheets. A pile of paper towels was placed on top of this stack, weighted down evenly and left to transfer overnight. The following morning the nylon membrane was removed, washed in 2× SSC for 2 min then placed on absorbent paper to partially dry. As soon as streaking appeared on the drying membrane, a CEX-800 UV cross-linker (Ultralum, Inc.) was used at 120 mJ/cm^2 to cross-link DNA to the membrane.

2.8.2 DIG-dUTP labelling of DNA probe

DNA (1 µg) was labelled for each probe using the DIG High Prime DNA Labelling and Detection Starter Kit (Roche), as per the manufacturer's instructions, resulting in approx. 20 µL of 100 ng µL^{-1} probe.

2.8.3 Probe hybridisation

Hybridisation of the DIG-labelled probe was performed using the DIG High Prime DNA Labelling and Detection Starter Kit (Roche), as per the manufacturer's instructions. Overnight hybridisation incubations were done in PYREX™ hybridisation tubes (Fisher Scientific) at 50°C with slow rotation.

2.8.4 Colourimetric development of Southern blots

Development of Southern blots was peformed using the DIG High Prime DNA Labelling and Detection Starter Kit (Roche), as per the manufacturer's instructions.

2.9 Transformation techniques

2.9.1 Preparation of fungal protoplasts (Yelton et al., 1985)

Liquid cultures of *E. festucae* or *P. paxilli* (2.4.2.1; 2.4.3.1) were incubated for 4 d at 22°C with shaking of 150 RPM or 30 h at 28°C with shaking of 200 RPM, respectively. Mycelia were harvested (2.4.4) from these cultures and washed three times with sterile water, followed by a final wash with OM buffer (1.2 M $MgSO_4$, 10 mM $NaHPO_4$, pH to 5.8 using NaH_2PO_4). Filter-sterilized lysing enzymes from *Trichoderma harzianum* (40 mL of 10 mg mL^{-1} stock in OM buffer; Sigma-Aldrich) was added to mycelia and incubated overnight at 22°C for *E. festucae* or 30°C for *P. paxilli* with shaking at 80 RPM. Samples were examined by microscopy to confirm protoplast formation and then filtered through a sterile coarse filter. The filtrate was

divided into 15 mL universal tubes (5 mL per tube) and overlaid with 2 mL ST buffer per tube (0.6 M sorbitol, 100 mM Tris, pH 8.0). Tubes were spun at 2375 × g for 5 min at 4°C, following which the protoplasts located at the interface between the lysing enzyme solution and ST buffer were removed by pipetting off the top layer with a 1 mL filter tip then placed in a fresh 15 mL universal tube. A 5 mL aliquot of STC buffer (1 M sorbitol, 50 mM Tris, 50 mM $CaCl_2$, pH 8.0) was added, mixed gently, and then spun at 2375 × g. This step was then repeated 3 times. Samples were pooled after each iteration until the protoplasts were contained within a single tube. Protoplasts were resuspended in 500 µL STC buffer, and a sample was taken for concentration estimation using a haemocytometer, with dilution in STC buffer as necessary. Aliquots (80 µL) of the protoplast suspension were then distributed into 1.5 mL microtubes and 20 µL of 40% (v/v) polyethylene glycol was added to each tube. The protoplasts were stored at -80°C until required.

2.9.2 Transformation of fungal protoplasts (Oliver et al., 1987)

Protoplast aliquots (2.9.1) were thawed on ice and 2 µL of 50 mM filter-sterilised spermidine and 5 µL of 5 mg mL^{-1} filter-sterilised heparin was added to each tube and mixed by gentle agitation. A 30 µL solution containing 2-5 µg of sample DNA, positive control DNA or water (protoplast viability and negative controls) was added to each tube. All tubes were mixed gently by agitation, left on ice for 30 min and 900 µL of 40% (v/v) polyethylene glycol was then added to each tube, agitated gently and maintained on ice. 50 µL aliquots of sample-transformed protoplasts and 100 µL aliquots of control-transformed protoplasts [2 × negative control and 1 × positive control (undiluted), 2 × negative control and 1 × positive control diluted 1:100 in STC buffer (1 M sorbitol, 50 mM Tris, 50 mM $CaCl_2$, pH 8.0)] were pipetted onto individual pre-poured RG plates (2.4% (w/v) potato dextrose, 0.8 M sucrose, 1.5% (w/v) agar, pH 6.5; 16 mL per plate). 3.5 mL of molten RG media (as above but with 0.8% (w/v) agar) previously equilibrated to 50°C was then pipetted onto these plates and mixed quickly but thoroughly with the protoplasts before setting, then incubated overnight at 22°C. The following day all plates were overlaid with 5 mL of molten RG media (0.8% agar equilibrated to 50°C), with the overlay of all plates except the protoplast viability controls containing sufficient quantity of the

selected antibiotic to achieve the desired concentration (2.4.2.3; 2.4.3.3) across the volume of the entire plate (approx. 24.5mL). Plates were left to grow for 7 days for *P. paxilli* or 14 days for *E. festucae*, or until colonies began emerging through the overlay. The colonies (approx. 200 for *E. festucae*) were transferred to plates containing solid media (2.3.1; 2.3.3) with appropriate antibiotic concentration for selection in the species being analysed (2.4.2.3; 2.4.3.3). Nuclear purification was achieved through single spore isolation for *P. paxilli* (2.4.3.5), or through 3× subculturing rounds for *E. festucae*.

2.9.3 Transformation of chemically-competent *E. coli* cells

Aliquots (50 µL) of DH5α chemically competent *E. coli* cells (Invitrogen) were thawed on ice, briefly centrifuged, and 1-5 µL of plasmid solution, ligation reaction or Gibson assembly reaction, was added and mixed through gentle agitation. Cells were incubated on ice for 30 min, heat-shocked at 42°C for 30 s, then placed back on ice for 5 min. SOC medium (2.3.4) (250 µL at 37°C) was added, and the samples were incubated for 1 h at 37°C while shaking at 200 RPM. Aliquots (30 µL and 270 µL) from each transformation were spread across separate LBA plates (2.3.2) using a sterilised glass spreader, and the plates were incubated at 37°C overnight to allow colonies to grow.

2.9.4 Transformation of electrocompetent *E. coli* cells

Aliquots (50 µL) of DH5α-E electrocompetent *E. coli* cells (Invitrogen) were thawed on ice. 1 µL of plasmid solution was added to each tube, mixed by agaitation and incubated on ice for 1 min. Cells were transferred to a cold 2 mm cuvette and electroporated once in a Gene Pulser (Bio-Rad) using settings of 25 mF, 2.5 kV and 200 Ω. Cells were suspended in 1mL of pre-warmed (37°C) SOC medium (2.3.4) and incubated at 37°C for 1 h without shaking. Aliquots were plated onto solid media (2.4.1.2) with appropriate antibiotic concentration for selection (2.4.1.3) and incubated overnight at 37°C to allow colonies to grow.

2.10 Plant methods

2.10.1 Plant inoculation (Latch and Christensen, 1985)

Perennial ryegrass seeds were surface sterilized by soaking in 50% (v/v) H_2SO_4 for 30 min, rinsed in tap water 3x, the soaking the seeds in 50% (v/v) chlorine bleach for 30 min, rinsing in sterile water 3x, then air dried on sterile filter paper in a laminar flow cabinet. Seeds were germinated on 3% (w/v) water agar in square 120 mm × 120 mm petri dishes, with 14 seeds placed on a line offset by 4 cm from one side of the petri dish. The plates were placed on their side (seeds at the bottom) in the dark for 7 d at 22°C in such a way that the emerging shoots could grow directly upwards. Using a sterile scalpel and dissecting microscope in a laminar flow cabinet, a shallow 2-3 mm longitudinal slit was cut into the 7 d old seedlings between the mesocotyl and coleoptile, and a small piece of freshly grown fungal mycelia was inserted into the cut. Seedlings were returned to the dark at 22°C for 7 d, checked for contamination after 2 d, then transferred to the light and incubated for another 7 d at 22°C. Seedlings were planted into root trainers containing potting mix and grown under standard conditions (2.10.2).

2.10.2 Standard plant growth conditions

Plants were grown under artificial lighting [diurnal cycle of 16 h light (approx. 100 µE m^{-2} s^{-1}) and 8 h dark] with temperature control at 19°C, though actual temperature for the plants was approx. 22°C due to overhead lighting. Plants were grown in root trainers with potting mix free from fungicide (AgResearch). Root trainers were placed in groups of 28 within larger trays, and watered with 800 mL of tap water twice a week, or as required.

2.10.3 Endophyte infection status determination by immunoblot

A single tiller from each plant to be tested for endophyte infection was cut transversely from as close to the base as possible using a scalpel. Dead leaf sheaths were then removed and the exposed cut was pressed to a nitrocellulose membrane (NCM). The NCM was immersed into freshly made blocking solution [20 mM Tris, 50 mM NaCl, 0.5% (w/v) non-fat milk powder] for 2 h at room temperature, then decanted off and replaced by 5 mL of blocking solution mixed with 5 µL of polyclonal

rabbit antibody raised against homogenised mycelium of *Epichloë festucae* var. *lolii* (Christensen et al., 1993; scaled up as necessary) and incubated overnight at 4°C with 15 RPM shaking. This solution was decanted off the following morning and the NCM washed 3 times in fresh blocking solution. Fresh blocking solution (5 mL) with 2.5 µL of secondary antibody (goat anti-rabbit antibody with an alkaline phosphatase conjugate, Sigma; scaled up as necessary) was added to the NCM and incubated on a shaker at room temperature for 2 h before decanting off and washing the NCM three times in blocking solution. Developing solution was prepared using SIGMAFAST™ Fast Red TR/Naphthol AS-MX tablets, as per the manufacturer's instructions, added to the NCM, and incubated on a shaker at room temperature for 15 min before decanting off and washing the NCM in water. The NCM was then dried on blotting paper and the results visualised.

2.10.4 Endophyte isolation from infected plant material

Pseudostem sections (4 cm) were removed near the base of infected plants using a scalpel. These were placed into a 15 mL universal tube (one per plant sampled) to which 95% (v/v) ethanol was added, and then poured off after 10 s. Chlorine bleach (10% v/v) was added to the tube and incubated at room temperature for 5 min. The bleach was decanted off, and the pseudostem sections were washed with sterile water. The pseudostem tissue was placed onto sterile filter paper and cut into 2 mm sections using a sterilised scalpel blade. Sterile forceps were used to separate these sections into individual layers, and single pieces of this material were placed onto petri dishes with PDA (2.3.3) containing 5 mg L^{-1} oxytetracycline. The samples were then incubated at 22°C until endophyte colonies grew sufficiently to be isolated from the plant material.

2.11 Metabolite extraction and analysis

2.11.1 Peramine extraction method

The weight of lyophilised *P. paxilli* mycelia for extraction was recorded (target weight varied from 20-100 mg across experiments), then mycelia were placed into 2 mL screw-cap microtubes along with two 2 mm diameter stainless steel beads and shaken in a FastPrep® FP120 Cell Disrupter System (Thermo Savant) for 20 s at 4

m s^{-1}, or until mycelia were completely powdered. Extraction solvent [1mL of 50% (v/v) methanol with 2.064 ng mL^{-1} homoperamine nitrate (AgResearch Grasslands) internal standard] was added to each tube and incubated for 1 h at room temperature in the dark with end-over-end rotation at 20 RPM. Samples were centrifuged for 10 min at 17,000 × g and 500 µL of supernatant was aspirated by syringe and transferred to an amber 12×32 mm HPLC vial through a 0.22-µm-pore polyvinylidene difluoride PVDF syringe filter.

2.11.2 Analysis of peramine by LCMS (standard method)

Samples were analysed by Wade Mace (AgResearch) using a MSQ™ mass spectrometer (Thermo Scientific). Each sample (5 µL injection) was chromatographically separated on a Synergi Polar-RP 100×2 mm (2.5 µm) column (Phenomenex, Torrance, CA) using a linear gradient profile (where eluent A is aqueous 0.1% formic acid and eluent B is acetonitrile), with time 0 (T_0) at 5% B, T_9 at 40% B, T_{11} at 90% B, and T_{13} at 90% B, followed by equilibration to initial conditions over the following 8 min. Peramine was quantified by mass spectroscopy (using homoperamine as an internal standard) according to the parameters described by Rasmussen et al. (2012). Peramine is expected to have a retention time of 8.6 min with an MS^1 ion of 248.1 *m/z*, and homoperamine is expected to have a retention time of 9.9 min with an MS^1 ion of 262.1 *m/z*.

2.11.3 Analysis of peramine by LCMS (linear ion-trap method)

Samples were analysed by Wade Mace (AgResearch) using a LTQ XL™ linear ion trap mass spectrometer (Thermo Scientific). Each sample (10 µL injection) was chromatographically separated on a Synergi Polar-RP 100×2 mm (2.5 µm) column (Phenomenex, Torrance, CA) using a linear gradient profile (where eluent A is aqueous 0.1% formic acid and eluent B is acetonitrile), with time 0 (T_0) at 5% B, T_9 at 40% B, T_{11} at 90% B, and T_{13} at 90% B, followed by equilibration to initial conditions over the following 8 min. Eluted compounds were detected using ESI ionisation (positive ion mode) with the MS^1 spectrum collected from 200-400 *m/z*. Additionally, two further MS^2 experiments were collected during the analysis to allow sensitive quantitation of peramine; 248.1 *m/z* MS^2 spectra as a sensitive detection of peramine, 262.1 *m/z* MS^2 spectra as a sensitive detection of homoperamine.

2.12 Microscopy

2.12.1 Visualisation of endophyte *in planta* by aniline blue staining

Grass pseudostems for analysis were cut with a scalpel blade into 5 cm sections and dead tissue was removed. The outermost leaf sheath was removed, and a thin epidermal peel was removed from the inside surface of this sheath by making a shallow transverse cut with a scalpel then gently peeling the uppermost layer off using forceps. The peels were placed on a microscope slide with aniline blue stain [25% (v/v) 88% lactic acid, 50% (v/v) glycerol, 0.1% (w/v) aniline blue] and a coverslip was placed on top. The slide was gently heated over a Bunsen burner flame until the stain began boiling to remove air bubbles and aid staining. Slides were observed on a light microscope, with fungal hyphae staining blue.

2.12.2 Identification of fusion events by inverted fluorescence microscopy (Lukito et al., 2015)

Growth media was prepared by placing a sterile microscopy slide onto a 20 mL base of solid media [1.5% (w/v) water agar] and overlaying with an additional 5 mL of this media. *E. festucae* mycelium were inoculated next to the edge of this microscopy slide and incubated at 22°C for 5 days. The microscopy slide was excised from the plate, with care taken to maintain the integrity of the agar overlay and attached *E. festucae* hyphae. Hyphae were stained with 250 µg mL^{-1} Calcofluor white (Fluorescent Brighter 28; Sigma-Aldrich) in 50% (v/v) glycerol and imaged on an IX71 inverted fluorescence microscope (Olympus) using the U-MWU2 ultraviolet excitation cube (wideband).

2.12.3 Confocal laser-scanning fluorescence microscopy

Grass pseudostems for analysis were cut with a scalpel blade into 5 cm sections and dead tissue was removed. Leaf sheaths were separated and further cut into 1 cm transverse sections, and these sections cut in half longitudinally again. The plant material was placed into microtubes containing 95% (v/v) ethanol and incubated at 4°C overnight, or stored for up to 2 weeks. Storage of material for longer than 2 weeks is not recommended, as autofluorescence is increased. Plant material was transferred into 1.5 mL microtubes containing 1 mL 10% (w/v) KOH and

soaked for 3 h at room temperature, or overnight at 4°C. Samples were carefully washed three times with 1 mL PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4) then placed in 1.5 mL microtubes containing 1 mL of freshly prepared staining solution [0.02% (w/v) aniline blue, 0.02% (v/v) Tween20, 0.001% (w/v) Alexa Fluor® 488 WGA (Invitrogen); optional 0.002% (w/v) propidium iodide added to visualise nuclei]. The tubes were incubated for 30 min at room temperature, including a 10 min vacuum infiltration step. Samples were placed onto microscope slides with PBS, covered with a coverslip and analysed using a Leica SP5 DM6000B confocal laser-scanning microscope at the Manawatu Microscopy and Imaging Centre. Images were captured using sequential excitation/capture windows of 405 nm/450-550 nm for aniline blue, 488 nm/500-550 nm for Alexa Fluor® 488 WGA and 520 nm/560-660 nm for propidium iodide. Image projections for z-sections were generated with ImageJ software using the maximum intensity method.

2.13 Bioinfomatic techniques

2.13.1 Multiple sequence alignments

Alignment of DNA sequences was performed with the ClustalW algorithm (Larkin et al., 2007) provided with MacVector 14.5 (MacVector, Inc.) using default parameters. Alignment of protein sequences was performed with the Muscle algorithm (Edgar, 2004) provided with MacVector 14.5 using default parameters.

2.13.2 Phylogenetic reconstruction of multiple sequence alignments

Phylogenetic reconstruction of DNA multiple sequence alignments (2.13.1) was performed using Mega 5.1 (Tamura et al., 2011) via the maximum likelihood method using the gamma-distributed (5 categories) Tamura three-parameter nucleotide substitution model (Tamura, 1992) and the subtree-pruning-regrafting (level 3) heuristic method on all sites and codons. Phylogenetic reconstruction of protein multiple sequence alignments (2.13.1) was performed using Mega 5.1 (Tamura et al., 2011) via the maximum likelihood method using default parameters. The bootstrap method with 1,000 repetitions was used to test all phylogenies.

2.14 Nucleotide sequence resources

Many of the *Epichloë* and Clavicipitaceaen genome sequences utilised are available on the Clavicipitaceae genome database (http://www.endophyte.uky.edu/) provided and maintained by Prof. Chris Schardl, University of Kentucky. Prof. Chris Schadl also kindly provided access to requested sequences from unpublished genomic and RNAseq data. Requested sequences from the AR1 and AR37 genomes were kindly provided by Richard Johnson (AgResearch).

3 Disparate independent genetic events disrupt the secondary metabolism gene *perA* in certain symbiotic *Epichloë* species

MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION TO DOCTORAL THESIS CONTAINING PUBLICATIONS

(To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis)

We, the candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: Daniel Berry

Name/Title of Principal Supervisor: Barry Scott

Name of Published Research Output and full reference:

Berry, D., Takach, J.E., Schardl, C.L., Charlton, N.D., Scott, B., and Young, C.A. (2015). Disparate independent genetic events disrupt the secondary metabolism gene perA in certain symbiotic Epichloë species. Appl Environ Microbiol 21, 2798-2807.

In which Chapter is the Published Work: 3

Please indicate either:

- The percentage of the Published Work that was contributed by the candidate: 80

  and / or

- Describe the contribution that the candidate has made to the Published Work:

  Generation, interpretation and analysis of all original experimental data except for experimental determination of peramine concentrations and the data presented in Fig. 3. Wrote publication text (with editing by Carolyn Young) and prepared all figures except Fig. 1 and Fig. 3.

Daniel Berry — Candidate's Signature

27/3/16 — Date

Barry Scott — Principal Supervisor's signature

27/3/16 — Date

Disparate Independent Genetic Events Disrupt the Secondary Metabolism Gene *perA* in Certain Symbiotic *Epichloë* Species

Daniel Berry,[a] Johanna E. Takach,[b] Christopher L. Schardl,[c] Nikki D. Charlton,[b] Barry Scott,[a] Carolyn A. Young[b]

Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand[a]; The Samuel Roberts Noble Foundation, Ardmore, Oklahoma, USA[b]; Department of Plant Pathology, University of Kentucky, Lexington, Kentucky, USA[c]

Peramine is an insect-feeding deterrent produced by *Epichloë* species in symbiotic association with C_3 grasses. The *perA* gene responsible for peramine synthesis encodes a two-module nonribosomal peptide synthetase. Alleles of *perA* are found in most *Epichloë* species; however, peramine is not produced by many *perA*-containing *Epichloë* isolates. The genetic basis of these peramine-negative chemotypes is often unknown. Using PCR and DNA sequencing, we analyzed the *perA* genes from 72 *Epichloë* isolates and identified causative mutations of *perA* null alleles. We found nonfunctional *perA*-ΔR* alleles, which contain a transposon-associated deletion of the *perA* region encoding the C-terminal reductase domain, are widespread within the *Epichloë* genus and represent a prevalent mutation found in nonhybrid species. Disparate phylogenies of adjacent A2 and T2 domains indicated that the deletion of the reductase domain (R*) likely occurred once and early in the evolution of the genus, and subsequently there have been several recombinations between those domains. A number of novel point, deletion, and insertion mutations responsible for abolishing peramine production in full-length *perA* alleles were also identified. The regions encoding the first and second adenylation domains (A1 and A2, respectively) were common sites for such mutations. Using this information, a method was developed to predict peramine chemotypes by combining PCR product size polymorphism analysis with sequencing of the *perA* adenylation domains.

Fungal secondary metabolites are a diverse group of important but often nonessential organic compounds with a wide range of properties that are likely to be advantageous for the producing organism or in some cases essential for pathogenicity or developmental stages (1–3). These low-molecular-weight compounds tend to only be produced under certain environmental or growth conditions. The biosynthetic pathways for production of any particular class of secondary metabolites are common to many fungi, but production of a specific secondary metabolite is often unique to a small phylogenetic group of species (4). *Epichloë* species are fungal endophytes of C_3 grasses that are known to produce several bioactive alkaloids that provide bioprotective properties to the host plant (5). These secondary metabolites include the indole-diterpenes, ergot alkaloids, lolines, and peramine (Fig. 1) (6, 7). The indole-diterpene lolitrem B and ergot alkaloid ergovaline have significant detrimental effects on the health and production of stock animals that graze infected pastures (7, 8). The lolines are insecticidal (9), and peramine is a potent deterrent of feeding by insects, including the agriculturally important invertebrate pest *Listronotus bonariensis* (Argentine stem weevil) (10–12).

Peramine synthesis is catalyzed by the two-module nonribosomal peptide synthetase (NRPS), peramine synthetase (PerA), encoded by the 8.3-kb gene *perA* (12). The first module of PerA contains an adenylation (A1) domain responsible for selection and activation of the proposed substrate amino acid 1-pyrroline-5-carboxylate and a thiolation (T1) domain that bonds this substrate as a thioester via a 4′-phosphopantetheine (4′PPT) linker. The second module contains adenylation (A2) and thiolation (T2) domains for selection, activation, and thiolation of the substrate proposed to be arginine. The second module also contains a methylation (M) domain proposed to N-methylate the alpha-amine of the arginine moiety, a condensation (C) domain that catalyzes peptide bond formation, and a variant reductase domain (R*) (13) at the C terminus, proposed to be responsible for intramolecular cyclization and release of the dipeptide product.

The genus *Epichloë* (including former *Neotyphodium* spp.) consists of sexual nonhybrid species and asexual, nonpathogenic endophytes that are derived either directly from the sexual species or by hybridization of two or more *Epichloë* progenitors (14, 15). Hybrid *Epichloë* species contain duplicate or even triplicate copies of most genes due to inheritance of an allele from each progenitor. Alleles of *perA* are found in nearly all *Epichloë* species, with the notable exceptions of *Epichloë glyceriae* and *Epichloë gansuensis* (16), but *perA* null alleles are common. One such allele, first identified in the genome sequence of *Epichloë festucae* isolate E2368 (16), has a deletion of the region encoding the C-terminal R* domain of PerA. This deletion is associated with the insertion of the miniature inverted-repeat transposable element (MITE) designated 3m (17). However, there are many other cases of peramine-negative (per$^-$) isolates for which the genetic basis is unknown (18, 19).

Peramine production is an important trait when considering

Received 12 November 2014 **Accepted** 5 February 2015

Accepted manuscript posted online 13 February 2015

Citation Berry D, Takach JE, Schardl CL, Charlton ND, Scott B, Young CA. 2015. Disparate independent genetic events disrupt the secondary metabolism gene *perA* in certain symbiotic *Epichloë* species. Appl Environ Microbiol 81:2797–2807. doi:10.1128/AEM.03721-14.

Editor: D. Cullen

Address correspondence to Carolyn A. Young, cayoung@noble.org.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/AEM.03721-14.

doi:10.1128/AEM.03721-14

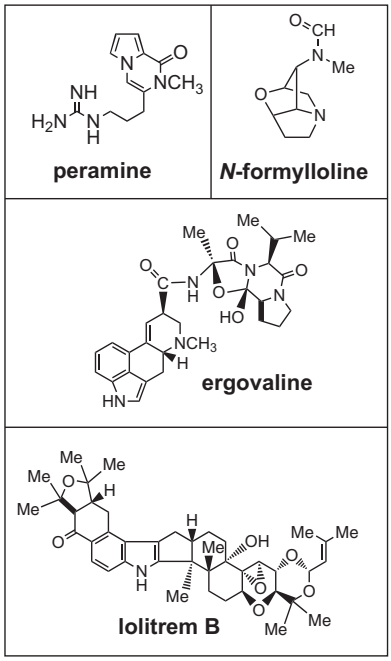


FIG 1 Chemical structures of alkaloid examples produced by *Epichloë* species.

endophyte strains for deployment in forage grasses and likely provides a selective advantage to endophyte-infected wild grasses. As such, diagnostic methods are useful to identify suitable endophyte isolates and associations for use in agriculture around the world. The objective of this study was to identify and characterize the mutations causing *perA* null alleles in a collection of hybrid and nonhybrid *Epichloë* species and strains. Using this information, we developed a PCR method to predict the peramine chemotype of endophytes from pure culture and in endophyte-infected plant material.

MATERIALS AND METHODS

Endophyte strains and growth conditions. Isolates of *Epichloë* species (see Table S1 in the supplemental material) were grown and maintained as previously described (20, 21). Endophyte-infected plant samples (see Table S2 in the supplemental material) were obtained from plant lines maintained under greenhouse conditions at 23°C during the day and 20°C at night, with 16 h of light starting at 0600; light intensity varied throughout the year, depending on the season.

Peramine analysis. Peramine was analyzed by AgResearch Grasslands (Palmerston North, New Zealand) from plant material using a modification of the method described by Rasmussen et al. (22). A 50-mg freeze-dried sample taken from endophyte-infected *Lolium perenne* whole tillers was extracted for 1 h with 1 ml of extraction solvent (50% [vol/vol] methanol) with 2.064 ng/ml homoperamine nitrate (AgResearch Grasslands) as an internal standard. The sample was then centrifuged for 5 min at 8,000 × g, and a 500-μl aliquot of the supernatant was transferred to an amber 12- by 32-mm high-performance liquid chromatography (HPLC) vial via a 0.22-μm-pore polyvinylidene difluoride (PVDF) syringe filter. Separation was achieved on a Synergi Polar-RP 100- by 2.00-mm (2.5-μm) column (Phenomenex, Torrance, CA) using a linear gradient profile (where eluent A is aqueous 0.1% formic acid and eluent B is acetonitrile), with time 0 (T_0) at 5% B, T_9 at 40% B, T_{11} at 90% B, and T_{13} at 90% B, followed by equilibration to initial conditions over the following 8 min. Peramine was quantified by mass spectroscopy (using homoperamine as an internal standard) according to the parameters described by Rasmussen et al. (22). Peramine is expected to have a retention time of 8.6 min with an MS1 ion of 248.1 *m/z*, and homoperamine is expected to have a retention time of 9.9 min with an MS1 ion of 262.1 *m/z*. A 5-μl injection volume gave a limit of detection for this technique of 0.1 μg/g for herbage.

Genomic DNA isolation. Genomic DNA was isolated from freeze-dried mycelium of *Epichloë* species using the ZR Fungal/Bacterial DNA MiniPrep kit (Zymo Research, Irvine, CA) as per the manufacturer's instructions. Total plant DNA (including endophyte) was extracted using the MagAttract 96 DNA plant core kit (Qiagen, Inc., Valencia, CA) as per the manufacturer's instructions.

Primer design. Primers for PCR amplification and sequencing (see Table S3 in the supplemental material) were designed using a multiple-sequence alignment of all available *perA* and flanking gene sequences from the *Epichloë* genome database (www.endophyte.uky.edu) (16), which included 10 strains from seven species. Primers were designed to maximize conservation of the target binding sequence between species. Primers for sequencing specific alleles from hybrid species were designed to contain at least two single nucleotide polymorphisms (SNPs) specific to each allele, with one of these SNPs located at the 3′ terminus wherever possible.

PCR amplification and product purification. Genomic DNA templates were amplified using GoTaq DNA polymerase (Promega, Madison, WI) under the conditions described by Takach et al. (18). PCR products for sequencing were purified using the QIAquick PCR purification kit (Qiagen). Where insufficient PCR product was produced for direct sequencing, a second PCR using a 10^2- or 10^3-fold dilution of the initial reaction was used as a template to increase PCR product concentration; this was often necessary when amplifying the *perA* gene directly from endophyte-infected plant material.

Sequencing of *perA*. Three overlapping DNA fragments, *perA*-1, *perA*-2, and *perA*-3, or *perA*-3ΔR*, covering the whole *perA* or *perA*-ΔR* gene, was amplified using primer sets defined in Table S3 in the supplemental material and sequenced with BigDye chemistry v3.1 (Applied Biosystems, Foster City, CA) using an Applied Biosystems 3730 DNA analyzer. Sequences from nonhybrid isolates were assembled using MacVector 12.6 with Assembler (MacVector, Inc., Cary, NC), with further sequencing completed using isolate-specific primers where required. Sequences from hybrid isolates were similarly assembled, but this assembly was then used as a reference to design allele-specific primers based on polymorphic regions to sequence each *perA* allele.

Phylogenetic reconstruction. A2 domain DNA sequences from 39 *perA* alleles (1,785 bp in length from positions 3575 to 5358 for *perA* from *E. festucae* Fl1) were aligned using ClustalW (23) and manually edited where necessary with MacVector 12.6. DNA sequences spanning from the middle of the T2 domain until the *perA*-ΔR* truncation location were similarly aligned from 38 *perA* alleles. These alignments were analyzed using Mega 5.1 (24) via the maximum likelihood method using the gamma-distributed (5 categories) Tamura three-parameter nucleotide substitution model (25) and the subtree-pruning-regrafting (level 3) heuristic method on all sites and codons. The bootstrap method with 1,000 repetitions was used to test the phylogeny.

Nucleotide sequence accession numbers. The GenBank accession numbers for *perA* sequences generated by this study are listed as follows: KP347845 to KP347877 and KP719965 to KP719973. Details about these and additional *perA* accession numbers from other studies (5, 22) are presented in Table S4 in the supplemental material.

RESULTS

Peramine chemotypes of *E. festucae* isolates. The distribution of peramine production within *E. festucae* was evaluated from herbage samples of *Lolium perenne* plants symbiotic with *E. festucae* isolate E189, Fg1, Fl1, Frc5, Frc7, Fr1, or Frr1. Of these associations, only the plants infected with Fl1, Frc7, or Frr1 contained peramine (Per^+) (Table 1). These data demonstrated that pe-

TABLE 1 Peramine concentrations of whole tillers from *Lolium perenne* infected by different *Epichloë festucae* isolates

E. festucae strain	Peramine concn (ppm)[a]	Gene feature
E189	ND[b]	*perA*-ΔR*
Fg1	ND	Deletion in A2 domain
Fl1	15–90[c]	*perA* functional
Frc5	ND	*perA*-ΔR*
Frc7	19.5	*perA* functional
Fr1	ND	*perA*-ΔR*
Frr1	137.2	*perA* functional

[a] Determined by combined liquid chromatography-mass spectroscopy. The limit of detection was 0.1 ppm, and the limit of quantification was 0.5 ppm.
[b] ND, not detected.
[c] Data from Tanaka et al. (12) and Young et al. (27).

ramine chemotypes can be highly variable and discontinuous, even between isolates of a single *Epichloë* species.

Analysis of *perA* across multiple *Epichloë* species. Genomic DNA extracted from mycelium of 34 different isolates spanning nine nonhybrid *Epichloë* species, including the *E. festucae* isolates mentioned above, was used in a PCR-based size polymorphism analysis to evaluate the presence and integrity of the *PER* locus. PCR primers were designed to amplify each *perA* domain in overlapping DNA fragments, as well as the conserved flanking genes *mfsA* and *qcrA*. The genes *mfsA* and *qcrA*, but not *perA*, were detected in *E. glyceriae* E2772 and *Epichloë elymi* E184 (Fig. 2). The majority of isolates (30/34) gave either a full complement of *perA* fragments or all fragments except the R*-domain fragment (no amplification with the primer pair perA3_3/perA3_R), indicating that these alleles likely lacked R* (which we designate *perA*-ΔR*). Although the regions encoding the A2 and C domains did not amplify from *Epichloë baconii* As6 and *Epichloë bromicola* E799, respectively, we know from sequencing and other PCR that these fragments are present (data not shown).

Alleles of *perA*-ΔR* were first observed in the genome sequence of *E. festucae* E2368, *E. festucae* var. lolii Lp14, and *Epichloë typhina* subsp. *poae* E5819 (16, 17). The region encoding the *perA* R* domain was also missing from both *Epichloë sylvatica* isolates tested and was discontinuously distributed within *E. baconii*, *E. bromicola*, *E. festucae*, and *E. typhina* (Fig. 2). An additional deletion was observed within the region encoding the T1 domain of *perA*-ΔR* from *E. sylvatica* isolates E354 and E503 (Fig. 2). Also detected was a deletion in the A2 domain fragment from the otherwise full-length *E. festucae* Fg1 *perA* allele (Fig. 2). The identification of *perA*-ΔR* in *E. festucae* E189, Frc5, and Fr1 and an A2 domain deletion in *E. festucae* Fg1 explains the observed per$^-$ chemotype of these isolates (Table 1).

Repetitive elements associated with *perA*-ΔR* alleles. Previous genome sequencing of *E. festucae* E2368, *E. festucae* var. lolii Lp14, and *E. typhina* subsp. *poae* E5819 indicated that deletion of the reductase domain was associated with repetitive elements, in particular the MITE 3m (16, 17). Additional representative isolates containing the *perA*-ΔR* alleles were analyzed by sequencing amplification products that span from the *perA* T2 domain to the adjacent gene, *qcrA* (Fig. 3A). This region failed to amplify from *E. baconii* As6, *E. baconii* E424, and *E. bromicola* E799. Comparison of the sequence data indicated variation of the repeats between isolates. In particular, different regions of MITE 3m were retained, and some isolates had the addition of MITE 25m (Fig. 3B). Also associated with the *perA*-ΔR* alleles is a unique 17-bp sequence located immediately downstream of the *perA* truncation (Fig. 3B). BLAST analysis using the 17-base sequence as a query revealed it is present only in the genome sequence of isolates containing *perA*-ΔR* and is not part of a repetitive element. The sequence was utilized as a primer (perA-17bp_R) (see Table S3 in the supplemental material) for PCR tests to determine if the 17-bp sequence was common to all *perA*-ΔR* alleles. DNA from all isolates identified as *perA*-ΔR* could be amplified with the perA-T2_F/perA-17bp_R primer set, confirming the association of the common 17-bp region with the deletion of the R* domain (Fig. 2 and 4). Thus, the perA-17bp_R primer allowed specific positive identification of the *perA*-ΔR* allele.

Analysis of *perA* from endophyte-infected plant material. *Epichloë* species for which no mycelium samples were readily available were evaluated directly from endophyte-infected plant samples. Total DNA was extracted from pseudostem or blade samples from 33 different *Epichloë*-infected plants spanning 13 grass species to evaluate the presence and integrity of *perA*. Of these plant symbionts, seven contained nonhybrid endophyte strains, 20 contained hybrid endophyte strains, and the hybrid status of endophytes in the remaining six samples was unknown. These DNA samples were used as the templates for PCR amplifications with the same primers described above for the mycelial genomic DNA templates. In most cases, PCR products were of significantly reduced intensity relative to those amplified from the mycelial genomic DNA samples (Fig. 4). This was expected because the endophyte DNA usually accounts for less than 2% of the total plant DNA (26, 27).

The majority of samples tested appeared to contain at least one intact *perA* gene. As previously shown, the tested *Epichloë* sp. FaTG-2 hybrid isolates NFe45079 and NFe45115 and the *Epichloë* sp. FaTG-3 hybrid isolate NFe1100 contained deletions in the regions encoding the A1 and A2 domains of the *perA* alleles known to be inherited from a *Lolium*-associated endophyte (LAE) progenitor (18, 28). *Epichloë cabralii* BlaTG-2 isolate NFe661, *Epichloë* sp. FaTG-2 G3 isolate NFe45115, and an isolate of *Epichloë uncinata* E167 have known per$^-$ chemotypes (18, 19, 29), yet each of these isolates appeared to contain at least one full-length *perA* allele (Fig. 4). The presence of full-length copies of *perA* suggested that small mutations within these gene copies likely generated *perA* null alleles.

The *perA*-ΔR* allele was identified in *E. typhina* OR10, *Epichloë siegelii* e915, and an undescribed endophyte, *Epichloë* sp. isolate e4768, from *Festuca versuta* (Fig. 4). The absence of the R*-domain product in *E. siegelii* e915, a two-parent hybrid, suggests that this strain contains two *perA*-ΔR* alleles. For *Epichloë* sp. strain e4768, the successful amplification of both the R*-domain and ΔR* deletion-specific PCR product indicates this isolate is a hybrid containing both the *perA* and *perA*-ΔR* alleles.

A draft genome sequence of *E. siegelii* e915 was used to explore the region flanking the two *perA*-ΔR* alleles (Fig. 3C). Annotation of *perA*-ΔR*, *mfsA*, *qcrA*, and repeat sequences that flank these genes revealed that the *perA*-ΔR* allele 1 was nearly identical to the arrangement found in *E. festucae* E2368 (Fig. 3B). Interestingly, the e915 *perA*-ΔR* allele 2, originating from the *E. bromicola* progenitor, was oriented toward *mfsA* rather than *qcrA*, indicating a gene inversion event has occurred. Although the common 17-bp region was still associated with this allele, there were no longer any downstream repetitive sequences. The

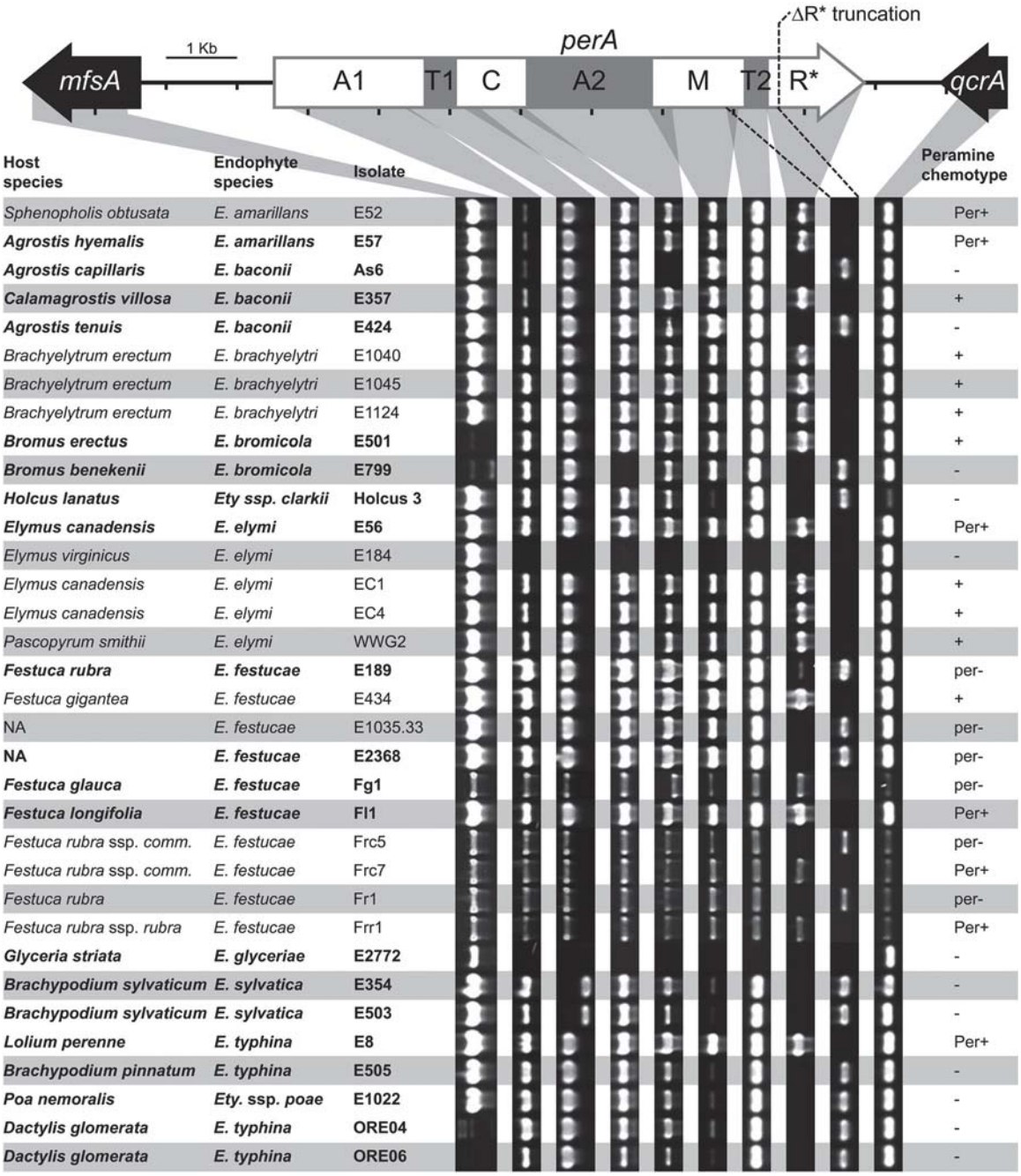


FIG 2 Analysis of *perA* integrity within nonhybrid *Epichloë* species. PCR-based size polymorphism analysis of *perA* from genomic DNA. The PCR products were produced using primers designed to the conserved sequence in or near each of the major *perA* domains and the two flanking genes *mfsA* and *qcrA* (see Table S3 in the supplemental material). Gray shading indicates the regions amplified. The dashed lines indicate the region that will amplify with primers perA-mid2-F2/perA_17bp_R only from isolates containing *perA*-ΔR* (Fig. 3). The regions encoding the adenylation (A), thiolation (T), condensation (C), methylation (M), and reductase* (R*) domains are shown within the *perA* gene map, with numbers indicating whether they are located in the first or second module. Isolates for which *perA* was subsequently sequenced or for which the *perA* sequence was already available (16, 33, 42) are shown in boldface. Low-intensity PCR products indicate the presence of SNPs in the primer target binding sequence. A host species labeled "NA" indicates the endophyte strain is the result of a controlled sexual cross. *Festuca rubra* subsp. *commutata* is abbreviated *Festuca rubra* ssp. *comm.* nt, not tested. Known peramine chemotypes are indicated as Per^+ (peramine producer) or per^- (peramine nonproducer) (5, 6, 16, 29, 43, 44). Peramine production was predicted for isolates with unknown peramine chemotypes based on the presence of the expected PCR products amplified from all *perA* domains and is indicated by + or −. Although not apparent in this screen, a *perA* remnant retaining the R* domain remains in E184 (39).

perA-ΔR* allele of *E. bromicola* isolate E799 contained a similar orientation (data not shown). Linkage between the contigs from e915 containing *perA*-ΔR* allele 2 and *qcrA* allele 2 cannot be determined from this sequence and were not able to be connected by PCR, likely due to the AT-rich repeat sequence that flanks *qcrA* (Fig. 3C).

Sequencing and characterization of *perA* variants. To determine why some isolates appeared to contain an intact *perA* gene

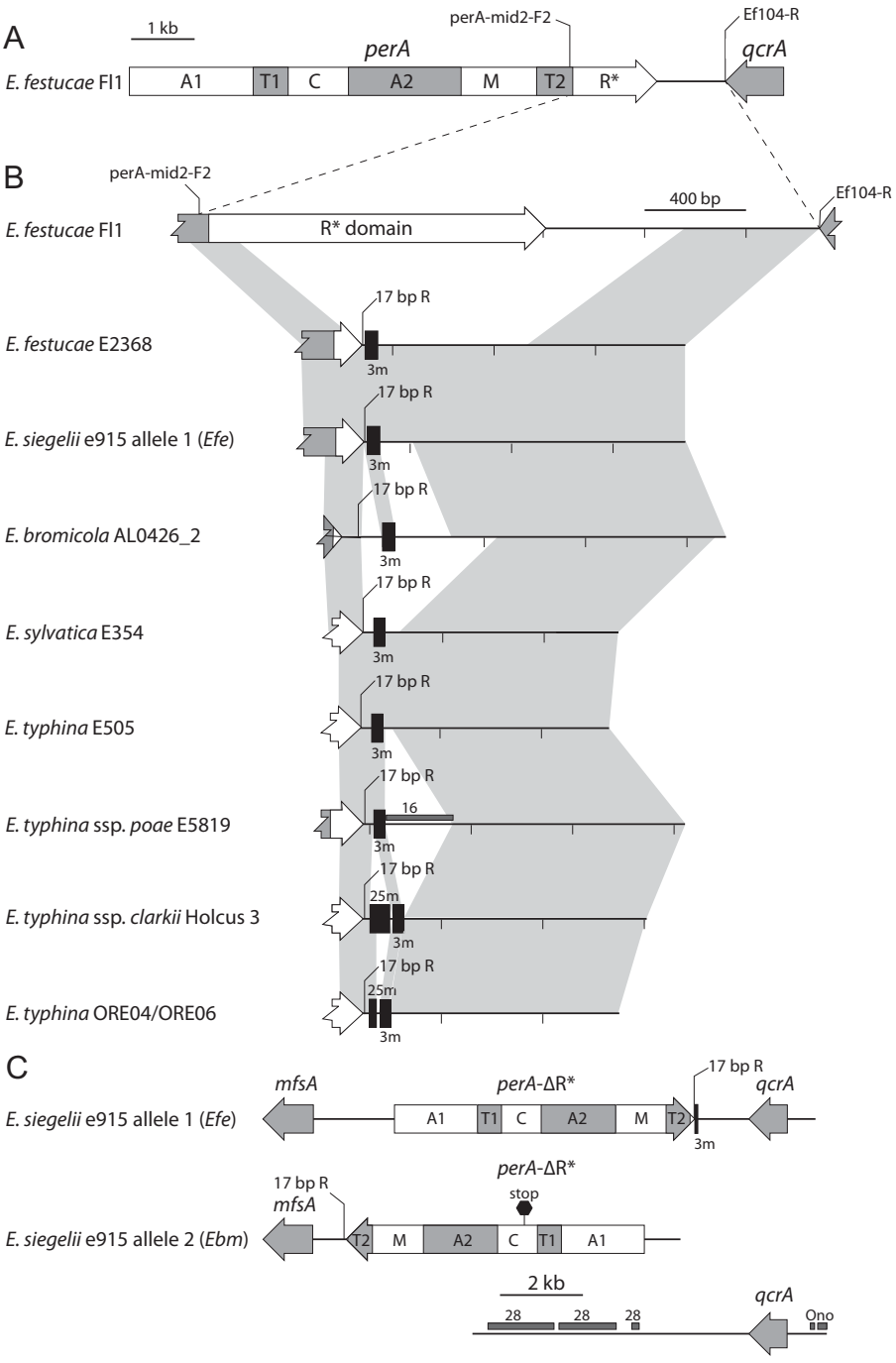
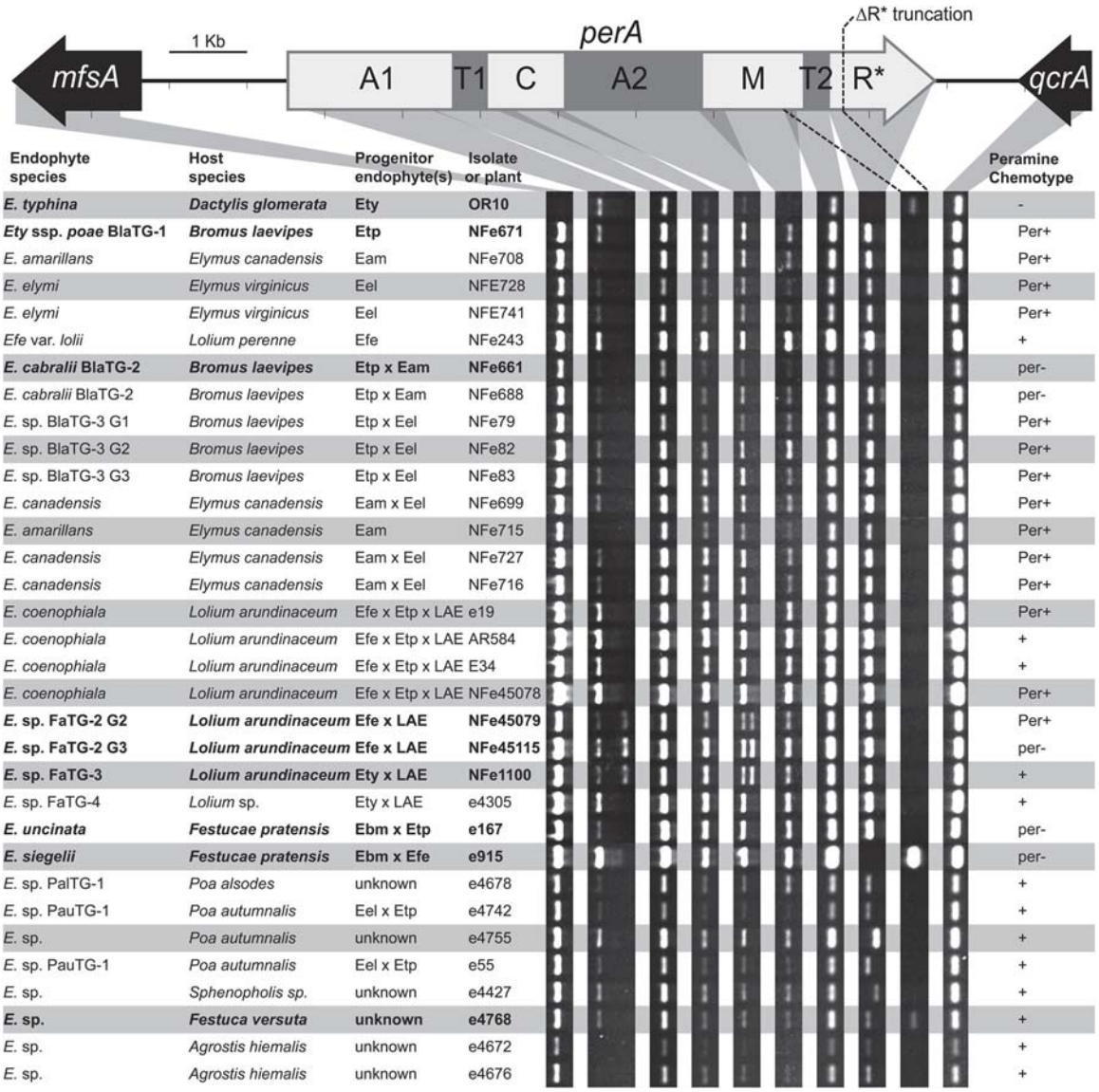


FIG 3 Analysis of *perA*-ΔR* downstream repeat sequences. Schematic representation of *Epichloë* isolates that lack the *perA*-R*domain. (A) Overview of the functional *perA* gene required for peramine production and the associated flanking gene *qcrA* from *E. festucae* Fl1. The domains of PerA are detailed in Fig. 2. (B) Schematic comparisons of regions from the *perA*-T2 domain to *qcrA*. The regions from *E. festucae* isolates Fl1 and E2368 and *E. typhina* subsp. *poae* E5819 were drawn from *perA* GenBank accession no. AB205145, JN640287, and JN640289, respectively, and *E. bromicola* AL0426_2 was generated from a genome sequence. The remaining examples were amplified with primer set perA-mid2-F2/Ef104-R using genomic DNA. Maps are arranged to illustrate synteny and do not necessarily suggest an evolutionary history. Syntenic regions (which may include small indels) between sequences are indicated by light gray polygons. Black vertical boxes indicate MITE 3m or 25m. Repeat sequence 16 (a putative retrotransposon) (16) is indicated by a dark gray horizontal box. The primer region for perA-17bp_R is shown in all *perA*-ΔR* sequences. (C) Schematic representation of the *PER* loci from a draft genome sequence of the hybrid species *E. siegelii* e915 demonstrating the inverted orientation of *perA*-ΔR* allele 2 relative to the flanking gene *mfsA*. Linkage between *perA*-ΔR* allele 2 and *qcrA* allele 2 is likely but cannot be proven due to the position of these genes on the ends of their respective contigs. The progenitor species from which each *E. siegelii* allele is derived is indicated in parentheses as "*Efe*" for *E. festucae* and "*Ebm*" for *E. bromicola*. Repeat sequences Ono and 28 (putative retrotransposons) (16) are indicated by dark gray horizontal boxes.



Endophyte species	Host species	Progenitor endophyte(s)	Isolate or plant	Peramine Chemotype
E. typhina	***Dactylis glomerata***	**Ety**	**OR10**	-
***Ety* ssp. *poae* BlaTG-1**	***Bromus laevipes***	**Etp**	**NFe671**	Per+
E. amarillans	*Elymus canadensis*	Eam	NFe708	Per+
E. elymi	*Elymus virginicus*	Eel	NFE728	Per+
E. elymi	*Elymus virginicus*	Eel	NFE741	Per+
Efe var. *lolii*	*Lolium perenne*	Efe	NFe243	+
***E. cabralii* BlaTG-2**	***Bromus laevipes***	**Etp x Eam**	**NFe661**	per-
E. cabralii BlaTG-2	*Bromus laevipes*	Etp x Eam	NFe688	per-
E. sp. BlaTG-3 G1	*Bromus laevipes*	Etp x Eel	NFe79	Per+
E. sp. BlaTG-3 G2	*Bromus laevipes*	Etp x Eel	NFe82	Per+
E. sp. BlaTG-3 G3	*Bromus laevipes*	Etp x Eel	NFe83	Per+
E. canadensis	*Elymus canadensis*	Eam x Eel	NFe699	Per+
E. amarillans	*Elymus canadensis*	Eam	NFe715	Per+
E. canadensis	*Elymus canadensis*	Eam x Eel	NFe727	Per+
E. canadensis	*Elymus canadensis*	Eam x Eel	NFe716	Per+
E. coenophiala	*Lolium arundinaceum*	Efe x Etp x LAE	e19	Per+
E. coenophiala	*Lolium arundinaceum*	Efe x Etp x LAE	AR584	+
E. coenophiala	*Lolium arundinaceum*	Efe x Etp x LAE	E34	+
E. coenophiala	*Lolium arundinaceum*	Efe x Etp x LAE	NFe45078	Per+
***E.* sp. FaTG-2 G2**	***Lolium arundinaceum***	**Efe x LAE**	**NFe45079**	Per+
***E.* sp. FaTG-2 G3**	***Lolium arundinaceum***	**Efe x LAE**	**NFe45115**	per-
***E.* sp. FaTG-3**	***Lolium arundinaceum***	**Ety x LAE**	**NFe1100**	+
E. sp. FaTG-4	*Lolium* sp.	Ety x LAE	e4305	+
E. uncinata	***Festucae pratensis***	**Ebm x Etp**	**e167**	per-
E. siegelii	***Festucae pratensis***	**Ebm x Efe**	**e915**	per-
E. sp. PalTG-1	*Poa alsodes*	unknown	e4678	+
E. sp. PauTG-1	*Poa autumnalis*	Eel x Etp	e4742	+
E. sp.	*Poa autumnalis*	unknown	e4755	+
E. sp. PauTG-1	*Poa autumnalis*	Eel x Etp	e55	+
E. sp.	*Sphenopholis sp.*	unknown	e4427	+
***E.* sp.**	***Festuca versuta***	**unknown**	**e4768**	+
E. sp.	*Agrostis hiemalis*	unknown	e4672	+
E. sp.	*Agrostis hiemalis*	unknown	e4676	+

FIG 4 Analysis of *perA* integrity from infected plant material. PCR analysis of *perA* from plants infected with *Epichloë* species. The PCR products were produced using primers designed to conserved sequence in or near each of the major *perA* domains and the two flanking genes *mfsA* and *qcrA* (see Table S3 in the supplemental material). Gray shading indicates the regions amplified. The dashed lines indicate the region that will amplify with primers perA-mid2-F2/perA_17bp_R only from isolates containing *perA*-ΔR* (Fig. 3). The PerA domains shown above the *perA* gene map are listed in Fig. 2. Isolates for which *perA* was subsequently sequenced are shown in boldface. Endophytes with unknown hybrid status are labeled "unknown." LAE, *Lolium*-associated endophyte; Eam, *E. amarillans*; Ebm, *E. bromicola*; Eel, *E. elymi*; Efe, *E. festucae*; Ety, *E. typhina*; Etp, *E. typhina* subsp. *poae*; nt, not tested. Known peramine chemotypes are indicated as Per^+ (peramine producer) or per^- (peramine nonproducer) (5, 6, 18, 19, 29, 43, 44). Peramine production was predicted for isolates with unknown peramine chemotypes based on the presence of expected PCR products amplified from all *perA* domains and are indicated by + or −.

but did not produce peramine, the *perA* alleles from 27 nonhybrid and seven hybrid isolates were sequenced or evaluated from genome sequences (Fig. 5; see Table S4 in the supplemental material). Of these 34 isolates, 11 were known to be Per^+, and 17 were known to be per^-; the peramine chemotypes of the remaining 6 isolates were unknown.

The identification of *perA*-ΔR* explained the per^- chemotype for 13 per^- isolates (Fig. 5C), and the presence of a deletion in the A2 domain in *E. festucae* Fg1 explained the per^- chemotype of this isolate (Fig. 5A). The remaining three per^- isolates, *E. cabralii* BlaTG-2 isolate NFe661, *E. uncinata* isolate e167, and *Epichloë* sp. FaTG-2 G3 isolate NFe45115, were hybrids (18, 19, 29). A 1-bp insertion causing a frameshift mutation was identified in *E. cabralii* BlaTG-2 NFe661 allele 1 (Fig. 5B), and an SNP that resulted in a nonsense mutation was identified in allele 2 (Fig. 5B). Analyses of the allele sequences from the hybrid isolate *E. uncinata* e167 identified independent frameshift mutations in alleles 1 and 2, generating *perA* null alleles (Fig. 5A). An SNP identified in the first adenylation domain of *Epichloë* sp. FaTG-2 G3 isolate NFe45115 allele 2 resulted in a nonsense mutation, while deletions are present in both A domains of allele 1 (Fig. 5A). These mutations explain the per^- chemotype of all three isolates.

The peramine chemotype was unknown for 6 of the 34 isolates from which *perA* was sequenced, so predictions were made by

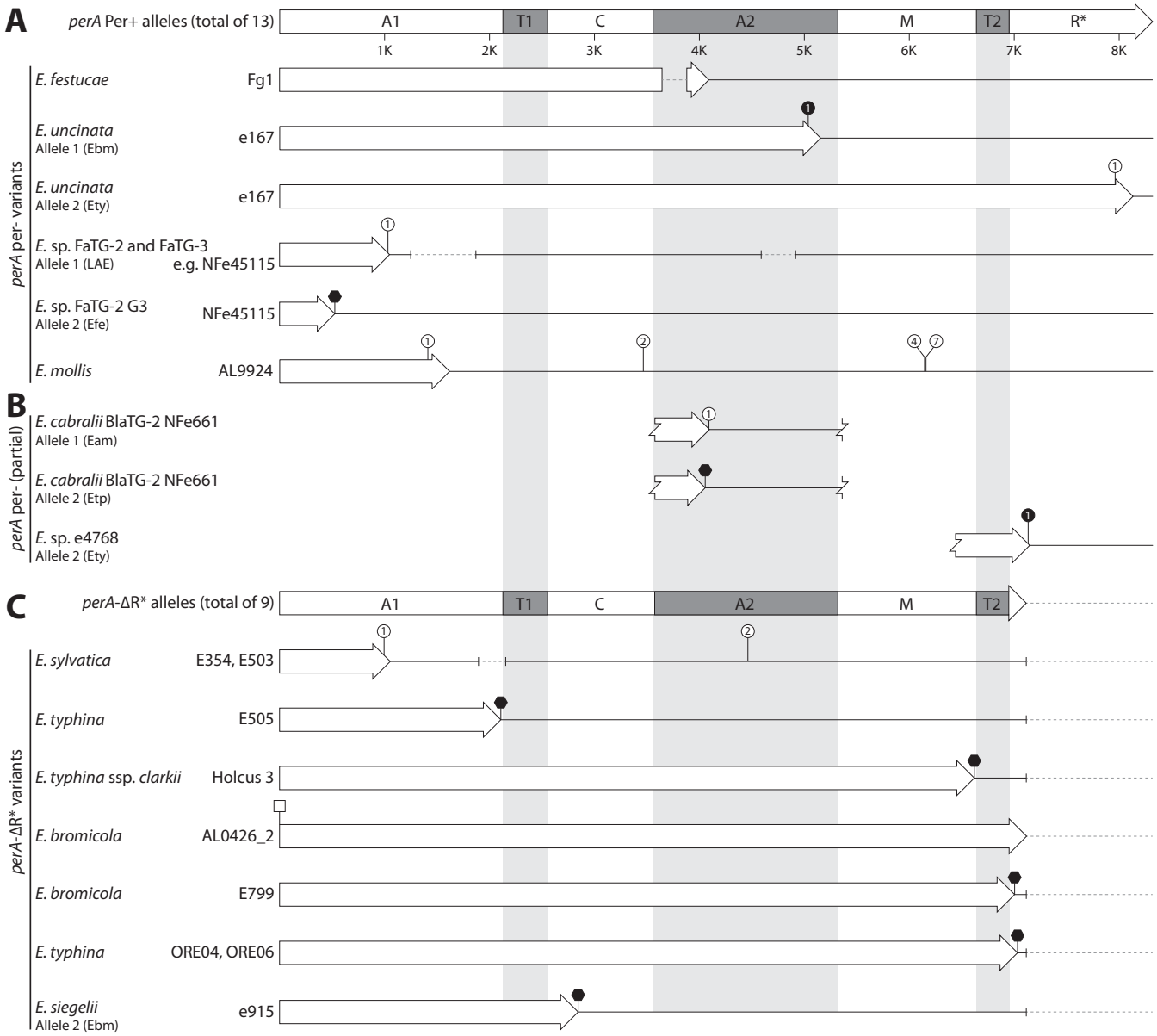


FIG 5 Genetic events disrupting *perA*. Gene maps of sequenced *perA* annotated with mutations likely to cause loss of function. (A) Gene maps of full-length *perA* sequences. The PerA domains shown on the functional *perA* gene map are listed in Fig. 2. The 13 Per$^+$ *perA* alleles represented by the first gene map are from *E. amarillans* E57 and E4668, *E. baconii* E1031, *E. brachyelytri* E4804, *E. bromicola* AL0434, E501, and E502, *E. elymi* E56, *E. festucae* Fl1, *E. typhina* E8, *E. typhina* subsp. *poae* BlaTG-1 isolate NFe671 (incomplete allele sequence), *Epichloë* sp. FaTG-2 G2 isolate NFe45079, and *Epichloë* sp. FaTG-3 isolate NFe1100. Gene maps of *perA* per$^-$ alleles detail causative mutations of full-length *perA* null alleles. (B) Partial gene maps of *perA* null alleles. (C) Gene maps of *perA*-ΔR* alleles. The 9 alleles represented by the initial gene map originate from *E. baconii* As6 and E424, *E. festucae* E189 and E2368 (identical alleles), *E. siegelii* e915 allele 1, *E. typhina* E1022 and OR10 (incomplete allele sequence), *E. typhina* subsp. *poae* E5819, and *Epichloë* sp. e4768 (incomplete allele sequence). Gene maps of other *perA*-ΔR* alleles detail changes that result in frameshift, nonsense, or large deletion mutations upstream of the conserved *perA*-ΔR* deletion site. Hollow arrows indicate the coding sequence for each allele, with the arrow ending at the first stop codon. Gene maps are annotated as follows. Frameshift-causing insertions or deletions are shown as white or black circles, respectively, with the number indicating how many nucleotides were inserted or deleted. SNPs that generate nonsense-mutation are shown as black hexagons. SNPs that disrupt the ATG start codon are shown as white squares. Large sequence deletions are shown as dashed lines, with the deletion shown relative to *perA* from *E. festucae* Fl1, and solid lines indicate conserved sequence located downstream of a premature stop codon. Eam, *E. amarillans*; Ebm, *E. bromicola*; Efe, *E. festucae*; Ety, *E. typhina*; Etp, *E. typhina* subsp. *poae*; LAE, *Lolium*-associated endophyte.

analyzing these sequence data. Of the six isolates, *E. baconii* E1031 and *E. bromicola* AL0434 were predicted to be Per$^+$ because they both contained full-length alleles with no nonsense, frameshift deletion, or insertion mutations. Of the four remaining isolates, *E. baconii* As6 and *E. bromicola* AL0426_2 were nonhybrid isolates containing *perA*-ΔR* alleles and were therefore predicted to be per$^-$. The *perA* allele from *Epichloë mollis* AL9924 contained multiple small (<10-bp) insertions causing frameshift mutations (Fig. 5A) and was therefore predicted to be per$^-$. Based on phylogenetic analysis, we judged the partial sequence from hybrid *Epichloë* sp. e4768 allele 2 to be derived from an *E. typhina* progenitor (Fig. 6B), and it is considered a *perA* null allele due to a

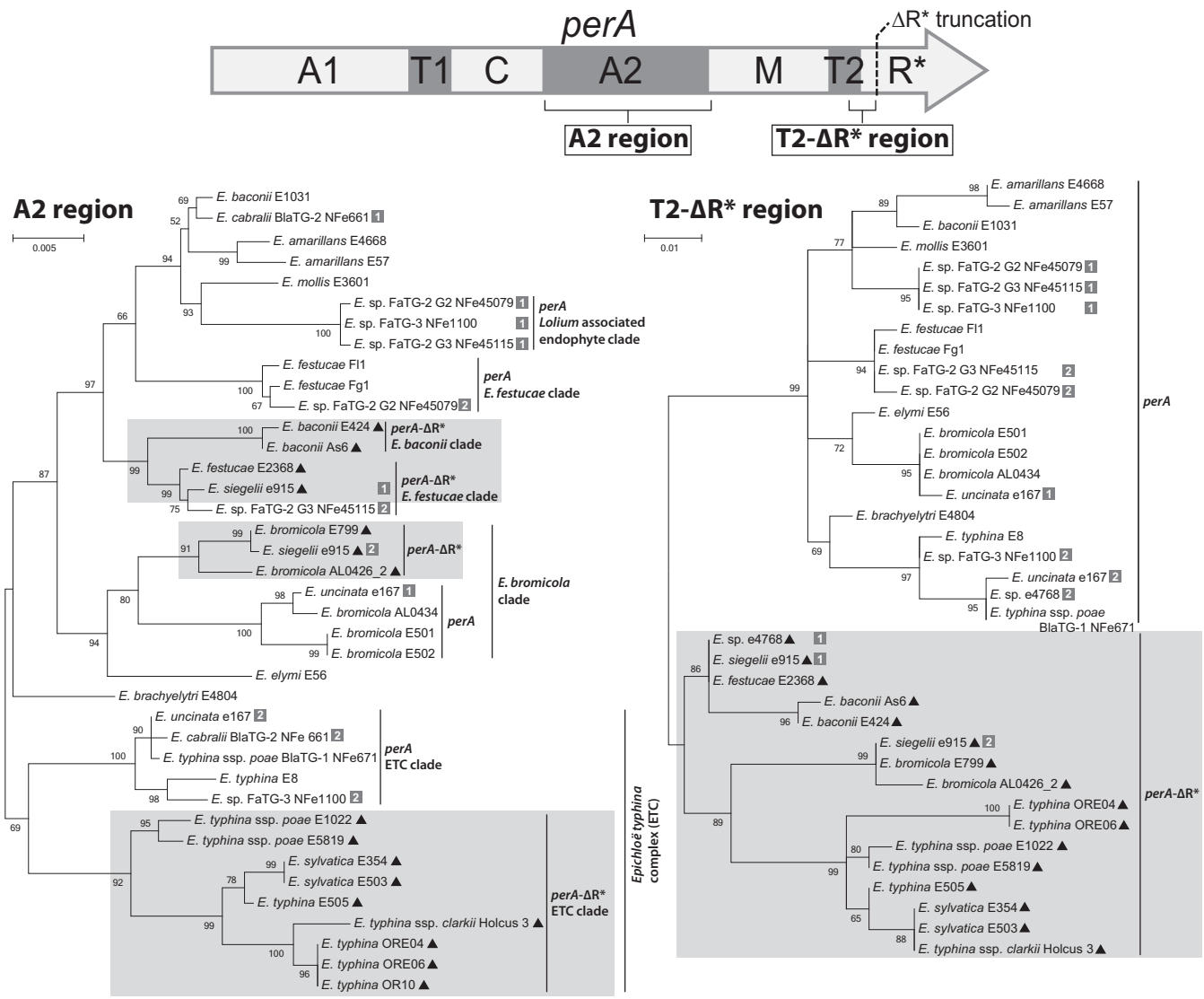


FIG 6 Phylogenetic tree of *perA* A2-domain DNA sequence. Shown are unrooted maximum likelihood phylogenetic trees generated from all available *perA* and *perA*-ΔR* A2-domain and *perA*-ΔR* sequences spanning from the middle of the T2 domain to the *perA*-ΔR* truncation location (T2-ΔR*), ending immediately prior to where homology between the *perA* and *perA*-ΔR* sequences stops. The regions used to generate each phylogenetic tree are shown relative to a *perA* gene map at the top of the figure, with PerA domain abbreviations detailed in the legend to Fig. 2. Black triangles next to isolate numbers identify the *perA*-ΔR* alleles. Numbers next to entries from hybrid species indicate which allele is represented.

1-bp deletion that results in a frameshift mutation (Fig. 5B). Isolate e4768 allele 1 was *perA*-ΔR* and was derived from an *E. festucae* E2368-like progenitor (Fig. 6B). Given that this isolate is a hybrid and has two null alleles, it is predicted to have a per$^-$ chemotype.

Analysis of the 18 sequenced *perA*-ΔR* alleles identified eight isolates that contained frameshift mutations, large deletions, and/or nonsense mutations in addition to the known 3m MITE-associated R*-domain deletion (Fig. 5C). A large deletion spanning the junction between the regions encoding the A1 and T1 domains and two small insertions were present in the alleles from *E. sylvatica* E354 and E503 (Fig. 5C), and nonsense mutations that should significantly truncate the translated protein were present in the *E. typhina* E505 and *E. typhina* subsp. *clarkii* Holcus 3 alleles. The *E. bromicola* AL0426_2 *perA*-ΔR* allele contained an SNP that disrupted the start codon, but a potential alternate ATG codon was located 189 bp downstream that does not truncate any conserved A-domain motifs. The nonsense mutations identified in the *perA*-ΔR* alleles from *E. bromicola* E799 and *E. typhina* ORE04 and ORE06 were located close to the existing *perA*-ΔR* truncation, and no other *perA* domains were affected (Fig. 5C). Both alleles from the hybrid isolate *E. siegelii* e915 were confirmed to be *perA*-ΔR*. A nonsense mutation that would result in truncation of the translated protein was identified in e915 allele 2 (Fig. 5C).

Phylogeny of *perA* A2- and T2-domain DNA sequences. Unrooted maximum likelihood phylogenetic trees were generated from the A2-domain DNA sequence of 39 *perA* and *perA*-ΔR* alleles, and the DNA sequence spanning from the middle of the T2 domain to the location of the *perA*-ΔR* truncation (T2-ΔR*) from 38 *perA* and *perA*-ΔR* alleles (Fig. 6). The A2-domain phylogeny revealed that all *E. typhina* complex-derived *perA*-ΔR* alleles grouped in a clade distinct from *E. typhina* complex *perA*

alleles (69% bootstrap support). The *E. bromicola*-derived *perA*-ΔR* alleles also grouped separately from the *E. bromicola perA* alleles (80% bootstrap support). The *perA*-ΔR* alleles derived from *E. festucae* and *E. baconii* isolates grouped separately from *perA* alleles from *Epichloë amarillans*, *E. baconii*, and *E. festucae* isolates (97% bootstrap support). In contrast, the phylogeny of the T2-ΔR* region showed all *perA*-ΔR* alleles grouped together separated from the *perA* alleles despite originating from multiple different species (Fig. 6). The A2 domain of FaTG-2 G3 NFe45115 *perA* allele 2 (derived from *E. festucae*) grouped with the *E. baconii*- and *E. festucae*-derived *perA*-ΔR* alleles, despite this allele still retaining the R* domain, as was previously observed by Takach et al. (18) (Fig. 6). In contrast the T2-ΔR* region from NFe45115 allele 2 grouped with related *perA* alleles (Fig. 6).

DISCUSSION

Peramine has been reported as the most commonly produced alkaloid by *Epichloë* species (66%), yet a discontinuous distribution is found within and between species (5, 6). In this study, we identify nonfunctional *perA* alleles from hybrid and nonhybrid *Epichloë* species. Although *E. glyceriae* and *E. gansuensis* have been previously shown to lack *perA* (16), we show the discontinuous distribution of peramine producers across *Epichloë* species is most frequently associated with mutations within *perA* that abolish peramine production. Analyses of apparently nonfunctional *perA* alleles show that each inactivating mutation is either isolate specific or is shared between closely related isolates of nonhybrids or closely related genomes in hybrids. Also, the DNA sequences encoding the first and second adenylation (A1 and A2) domains are common sites for such inactivating mutations (Fig. 5). These data indicate that independent mutation events have inactivated *perA* many times. The only exception to this rule is the identical R*-domain deletion found in all *perA*-ΔR* alleles. Identification of these inactivating *perA* mutations provides information to aid in prediction of peramine producers for isolates with unknown peramine chemotypes and genetic diagnosis for isolates known to be per$^-$.

In this study, the *PER* locus was evaluated using PCR to assess the integrity of each *perA* domain from 67 isolates representing at least 20 *Epichloë* species. We were able to distinguish the full-length *perA* alleles found in peramine producers and the *perA*-ΔR* alleles found in strains that are unable to produce peramine. Amplification of genomic DNA using a primer pair specific to *perA*-ΔR* (perA-T2_F and perA-17bp_R) identified 18 (27%) isolates missing the R* domain, of which two isolates, *E. siegelli* e915 and *Epichloë* sp. e4768, were hybrid species and the remaining were nonhybrids. However, the PCR analyses used to detect the presence of each domain did not reveal all *perA* mutants. Sequencing of *perA* from isolates unable to make peramine revealed frameshift and nonsense mutations predominantly within the regions encoding the A1 and A2 domains that would render *perA* nonfunctional. From these sequence data, we were able to explain the mutations responsible for the peramine-negative chemotype previously identified in *E. uncinata* e167, *Epichloë* sp. FaTG-2 G3 isolate NFe45115, and *E. cabralii* BlaTG-2 isolate NFe661 (18, 19, 29).

The isolate- or lineage-specific nature of mutations that have resulted in *perA* null alleles contrasts sharply with the taxonomic distribution of *perA*-ΔR*. The *perA*-ΔR* alleles are distributed widely within the *Epichloë* genus, occurring in a subset of isolates from each of the *E. baconii*, *E. bromicola*, *E. festucae*, and *E. typhina* complex (ETC) clades (Fig. 6). Given that the R* domain was deleted at identical sites within *perA* and there is high sequence conservation immediately downstream of the *perA*-ΔR* alleles, it is unlikely this deletion occurred more than once. In support of the possibility that a single event was responsible is the consistent association of *perA*-ΔR* with a downstream MITE 3m sequence and a unique 17-bp sequence containing an in-frame stop codon. Both of these features were absent from all *perA* alleles, so the MITE 3m insertion seems likely to have been involved in the deletion of the region encoding the R* domain.

The evolution of the *perA*-ΔR* alleles appears particularly complex, considering the disparate phylogenies of the regions encoding A2 and T2-ΔR* (Fig. 6). The T2-ΔR*phylogeny placed *perA* and *perA*-ΔR* into separate clades, each independently reflecting relationships of the *Epichloë* species. The T2-ΔR* phylogeny suggested transspecies polymorphism (TSP), whereby the corresponding sequences in *perA* and *perA*-ΔR* diverged early during, or even before, evolution of the genus *Epichloë*. This pattern is similar to evidence of TSP in other systems, such as vertebrate major histocompatibility loci (30) and fungal vegetative incompatibility loci (31).

In contrast to the T2 phylogeny, the phylogeny of the A2-encoding sequences consistently grouped *perA*-ΔR* alleles with *perA* of the same or closely related species, although in most species, the separation of the *perA*-ΔR* and *perA* subclades seemed deeply rooted in the species. The disparity between the A2 and T2-ΔR* phylogenies suggests multiple recombination events. What appears to be the most recent example affected *perA* allele 2 in FaTG-2 G3 isolate NFe45115. The region encoding the A2 domain of this *perA* allele groups with the *E. festucae* and *E. baconii perA*-ΔR* alleles, whereas the T2-ΔR* sequence groups with *E. festucae perA* alleles (Fig. 6). The fact that multiple species clades exhibit the disparate A2 and T2-ΔR* phylogenies suggests that a recombination hot spot exists between these two portions of *perA*.

The only two hybrid isolates containing *perA*-ΔR* alleles were *E. siegelii* isolate e915 and *Epichloë* sp. e4768. This is perhaps surprising given the wide distribution seen in the sexual isolates (44% of isolates tested in this study) (Fig. 2) and the number of hybrid species we tested (21 isolates representing 10 species) that contain *E. festucae* (7 isolates) and *E. typhina* (14 isolates) ancestral progenitors.

Previous studies of *Epichloë* alkaloid biosynthesis loci, such as the ergot alkaloid (*EAS*), indole-diterpene (*IDT*/*LTM*), and loline (*LOL*) gene clusters, have shown the presence or absence of pathway-specific genes to be the primary factor determining chemotype diversity of these alkaloids (16, 32, 33). The *IDT*/*LTM* and *EAS* loci are localized to dynamic subterminal regions of chromosomes (16), and both these and the *LOL* gene cluster are closely associated with a variety of transposable elements (16, 17, 34). These factors provide mechanisms through which genes from these clusters, or even an entire gene cluster, may be lost via recombination when selective pressure for a cluster is reduced. In contrast, with the exception of *perA*-ΔR* alleles, full-length *perA* alleles have not been found in association with transposable elements (16). In the absence of selective pressure, *perA* is likely to be retained longer than genes from the other secondary metabolite gene clusters, and this could explain the observed increase of *perA* inactivation by nonsense, frameshift, or deletion mutations rela-

tive to gene loss events common to the *EAS*, *IDT*/*LTM* and *LOL* gene clusters.

Diagnostic PCR utilizing markers developed from sequences of housekeeping and secondary metabolite biosynthetic genes is an effective approach to identify and quantify potential contamination from mycotoxin-producing fungi within foodstuffs for human and animal consumption (35, 36). For example, multiplex PCRs have been successfully used to simultaneously detect multiple fungal genera found in cereals that are likely to produce ochratoxins and trichothecenes (37). A quantitative PCR (qPCR) assay that detects polymorphisms within *TRI12* can identify different trichothecene genotypes within *Fusarium* species from field samples (38). Chemotype prediction using PCR to detect the presence of biosynthesis genes has also been very successful when evaluating *Epichloë* species for the ability to produce ergot alkaloids, indole-diterpenes, and lolines and provides insight into the bioactive potential of any given endophyte isolate (18, 19, 32, 39–41). In all of these approaches, the ability to directly analyze infected plant material by PCR provides rapid detection methods for a wide range of organisms and their biosynthetic potential. To determine whether an endophyte isolate is likely to produce peramine, we have refined the PCR approach described previously (18, 19, 39, 41) in order to identify the presence and integrity of all domains encoded by *perA*. In addition, sequence analysis of the regions encoding the A1 and A2 domains can be used to identify the most commonly found mutations. Using this pipeline, specific isolates with known and unknown peramine chemotypes were screened to identify *perA*-ΔR* alleles and other observable deletions, and sequence analysis was used to identify frameshift and nonsense mutations that would render *perA* nonfunctional. Although this method will not eliminate the need to evaluate peramine production, especially for determination of the levels of peramine produced by a given isolate, it does provide insight into the likelihood of peramine production. Evaluation of endophyte-infected plant germplasm for potential peramine producers as well as production of other bioactive alkaloids will help us understand the bioprotective potential of *Epichloë* species and facilitate investigation into the effects of different geographic and selective pressures on the evolution of this locus.

ACKNOWLEDGMENTS

We acknowledge Ginger A. Swoboda (The Samuel Roberts Noble Foundation) for technical assistance, Pierre-Yves Dupont (Massey University) for assistance with phylogeny reconstruction methodology, and Wade Mace (AgResearch) for measurement of peramine concentrations. We thank Adrian Leuchtmann (ETH Zurich) for access to *Epichloë bromicola* (AL0434 and AL0426_2) genome sequences.

This research was supported by a Massey University Ph.D. scholarship and a grant from the Royal Society of New Zealand Marsden Fund (contract MAU1002). Genome sequencing was supported by United States Department of Agriculture grants 2012-67013-19384 and 2010-34457-21269, National Institutes of Health grants R01GM086888 and 2 P20 RR-16481, and The Samuel Roberts Noble Foundation.

REFERENCES

1. **Keller NP, Turner G, Bennett JW.** 2005. Fungal secondary metabolism—from biochemistry to genomics. Nat Rev Microbiol **3:**937–947. http://dx.doi.org/10.1038/nrmicro1286.
2. **Forseth RR, Amaike S, Schwenk D, Affeldt KJ, Hoffmeister D, Schroeder FC, Keller NP.** 2013. Homologous NRPS-like gene clusters mediate redundant small-molecule biosynthesis in *Aspergillus flavus*. Angew Chem Int Ed Engl **52:**1590–1594. http://dx.doi.org/10.1002/anie.201207456.
3. **Scharf DH, Heinekamp T, Brakhage AA.** 2014. Human and plant fungal pathogens: the role of secondary metabolites. PLoS Pathog **10**(1)**:** e1003859. http://dx.doi.org/10.1371/journal.ppat.1003859.
4. **Kroken S, Glass NL, Taylor JW, Yoder OC, Turgeon BG.** 2003. Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. Proc Natl Acad Sci U S A **100:**15670–15675. http://dx.doi.org/10.1073/pnas.2532165100.
5. **Clay K, Schardl C.** 2002. Evolutionary origins and ecological consequences of endophyte symbiosis with grasses. Am Nat **160:**S99–S127. http://dx.doi.org/10.1086/342161.
6. **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. http://dx.doi.org/10.1007/BF00982100.
7. **Bush LP, Wilkinson HH, Schardl CL.** 1997. Bioprotective alkaloids of grass-fungal endophyte symbioses. Plant Physiol **114:**1–7.
8. **Bluett SJ, Thom ER, Clark DA, Macdonald KA, Minnee EMK.** 2005. Effects of perennial ryegrass infected with either AR1 or wild endophyte on dairy production in the Waikato. N Z J Agric Res **48:**197–212. http://dx.doi.org/10.1080/00288233.2005.9513650.
9. **Schardl CL, Grossman RB, Nagabhyru P, Faulkner JR, Mallik UP.** 2007. Loline alkaloids: currencies of mutualism. Phytochemistry **68:**980–996. http://dx.doi.org/10.1016/j.phytochem.2007.01.010.
10. **Rowan DD, Hunt MB, Gaynor DL.** 1986. Peramine, a novel insect feeding deterrent from ryegrass infected with the endophyte *Acremonium loliae*. J Chem Soc Chem Commun **1986:**935–936. http://dx.doi.org/10.1039/C39860000935.
11. **Siegel MR, Bush LP.** 1997. Toxin production in grass/endophyte associations, p 185–207. *In* Carroll G, Tudzynski P (ed), Plant relationships, vol 5. Springer, Berlin, Germany.
12. **Tanaka A, Tapper BA, Popay A, Parker EJ, Scott B.** 2005. A symbiosis expressed non-ribosomal peptide synthetase from a mutualistic fungal endophyte of perennial ryegrass confers protection to the symbiotum from insect herbivory. Mol Microbiol **57:**1036–1050. http://dx.doi.org/10.1111/j.1365-2958.2005.04747.x.
13. **Liu XY, Walsh CT.** 2009. Cyclopiazonic acid biosynthesis in *Aspergillus* sp.: characterization of a reductase-like R* domain in cyclopiazonate synthetase that forms and releases cyclo-acetoacetyl-L-tryptophan. Biochemistry (Wash) **48:**8746–8757. http://dx.doi.org/10.1021/bi901123r.
14. **Leuchtmann A, Bacon CW, Schardl CL, White JF, Jr, Tadych M.** 2014. Nomenclatural realignment of *Neotyphodium* species with genus *Epichloë*. Mycologia **106:**202–215. http://dx.doi.org/10.3852/13-251.
15. **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. http://dx.doi.org/10.1111/j.1365-294X.2004.02138.x.
16. **Schardl CL, Young CA, Hesse U, Amyotte SG, Andreeva K, Calie PJ, Fleetwood DJ, Haws DC, Moore N, Oeser B, Panaccione DG, Schweri KK, Voisey CR, Farman ML, Jaromczyk JW, Roe BA, O'Sullivan DM, Scott B, Tudzynski P, An Z, Arnaoudova EG, Bullock CT, Charlton ND, Chen L, Cox M, Dinkins RD, Florea S, Glenn AE, Gordon A, Gueldener U, Harris DR, Hollin W, Jaromczyk J, Johnson RD, Khan AK, Leistner E, Leuchtmann A, Li C, Liu J, Liu J, Liu M, Mace W, Machado C, Nagabhyru P, Pan J, Schmid J, Sugawara K, Steiner U, Takach JE, Tanaka E, Webb JS, Wilson EV, Wiseman JL, Yoshida R, Zeng Z.** 2013. Plant-symbiotic fungi as chemical engineers: multi-genome analysis of the Clavicipitaceae reveals dynamics of alkaloid loci. PLoS Genet **9**(2)**:**e1003323. http://dx.doi.org/10.1371/journal.pgen.1003323.
17. **Fleetwood DJ, Khan AK, Johnson RD, Young CA, Mittal S, Wrenn RE, Hesse U, Foster SJ, Schardl CL, Scott B.** 2011. Abundant degenerate miniature inverted-repeat transposable elements in genomes of epichloid fungal endophytes of grasses. Genome Biol Evol **3:**1253–1264. http://dx.doi.org/10.1093/gbe/evr098.
18. **Takach JE, Mittal S, Swoboda GA, Bright SK, Trammell MA, Hopkins AA, Young CA.** 2012. Genotypic and chemotypic diversity of *Neotyphodium* endophytes in tall fescue from Greece. Appl Environ Microbiol **78:**5501–5510. http://dx.doi.org/10.1128/AEM.01084-12.
19. **Charlton ND, Craven KD, Afkhami ME, Hall BA, Ghimire SR, Young CA.** 2014. Interspecific hybridization and bioactive alkaloid variation increases diversity in endophytic *Epichloë* species of *Bromus laevipes*. FEMS Microbiol Ecol **90:**276–289. http://dx.doi.org/10.1111/1574-6941.12393.
20. **Moon CD, Tapper BA, Scott B.** 1999. Identification of *Epichloë* endo-

phytes *in planta* by a microsatellite-based PCR fingerprinting assay with automated analysis. Appl Environ Microbiol **65:**1268–1279.
21. **Moon CD, Scott B, Schardl CL, Christensen MJ.** 2000. The evolutionary origins of *Epichloë* endophytes from annual ryegrasses. Mycologia **92:** 1103–1118. http://dx.doi.org/10.2307/3761478.
22. **Rasmussen S, Lane GA, Mace W, Parsons AJ, Fraser K, Xue H.** 2012. The use of genomics and metabolomics methods to quantify fungal endosymbionts and alkaloids in grasses. Methods Mol Biol **860:**213–226. http://dx.doi.org/10.1007/978-1-61779-594-7_14.
23. **Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG.** 2007. Clustal W and Clustal X version 2.0. Bioinformatics **23:**2947–2948. http://dx.doi.org/10.1093/bioinformatics/btm404.
24. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S.** 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol **28:**2731–2739. http://dx.doi.org/10.1093/molbev/msr121.
25. **Tamura K.** 1992. Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. Mol Biol Evol **9:**678–687.
26. **Panaccione DG, Johnson RD, Wang JH, 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 U S A **98:**12820–12825. http://dx.doi.org/10.1073/pnas.221198698.
27. **Young CA, Bryant MK, Christensen MJ, Tapper BA, Bryan GT, Scott B.** 2005. Molecular cloning and genetic analysis of a symbiosis-expressed gene cluster for lolitrem biosynthesis from a mutualistic endophyte of perennial ryegrass. Mol Genet Genomics **274:**13–29. http://dx.doi.org/10.1007/s00438-005-1130-0.
28. **Ekanayake PN, Rabinovich M, Guthridge KM, Spangenberg GC, Forster JW, Sawbridge TI.** 2013. Phylogenomics of fescue grass-derived fungal endophytes based on selected nuclear genes and the mitochondrial gene complement. BMC Evol Biol **13:**270. http://dx.doi.org/10.1186/1471-2148-13-270.
29. **Leuchtmann A, Schmidt D, Bush LP.** 2000. Different levels of protective alkaloids in grasses with stroma-forming and seed-transmitted *Epichloë/Neotyphodium* endophytes. J Chem Ecol **26:**1025–1036. http://dx.doi.org/10.1023/A:1005489032025.
30. **Cutrera AP, Lacey EA.** 2007. Trans-species polymorphism and evidence of selection on class II MHC loci in tuco-tucos (Rodentia: Ctenomyidae). Immunogenetics **59:**937–948. http://dx.doi.org/10.1007/s00251-007-0261-3.
31. **Muirhead CA, Glass NL, Slatkin M.** 2002. Multilocus self-recognition systems in fungi as a cause of trans-species polymorphism. Genetics **161:** 633–641.
32. **Young CA, Tapper BA, May K, Moon CD, Schardl CL, Scott B.** 2009. Indole-diterpene biosynthetic capability of *Epichloë* endophytes as predicted by *ltm* gene analysis. Appl Environ Microbiol **75:**2200–2211. http://dx.doi.org/10.1128/AEM.00953-08.
33. **Schardl CL, Young CA, Pan J, Florea S, Takach JE, Panaccione DG, Farman ML, Webb JS, Jaromczyk J, Charlton ND, Nagabhyru P, Chen L, Shi C, Leuchtmann A.** 2013. Currencies of mutualisms: sources of alkaloid genes in vertically transmitted epichloae. Toxins **5:**1064–1088. http://dx.doi.org/10.3390/toxins5061064.
34. **Young CA, Felitti S, Shields K, Spangenberg G, Johnson RD, Bryan GT, Saikia S, Scott B.** 2006. A complex gene cluster for indole-diterpene biosynthesis in the grass endophyte *Neotyphodium lolii*. Fungal Genet Biol **43:**679–693. http://dx.doi.org/10.1016/j.fgb.2006.04.004.
35. **Niessen L.** 2007. PCR-based diagnosis and quantification of mycotoxin producing fungi. Int J Food Microbiol **119:**38–46. http://dx.doi.org/10.1016/j.ijfoodmicro.2007.07.023.
36. **Gong L, Jiang Y, Chen F.** 25 September 2014. Molecular strategies for detection and quantification of mycotoxin-producing *Fusarium* species: a review. J Sci Food Agric http://dx.doi.org/10.1002/jsfa.6935.
37. **Vegi A, Wolf-Hall CE.** 2013. Multiplex real-time PCR method for detection and quantification of mycotoxigenic fungi belonging to three different genera. J Food Sci **78:**M70–M76. http://dx.doi.org/10.1111/j.1750-3841.2012.03008.x.
38. **Nielsen LK, Jensen JD, Rodriguez A, Jorgensen LN, Justesen AF.** 2012. *TRI12* based quantitative real-time PCR assays reveal the distribution of trichothecene genotypes of *F. graminearum* and *F. culmorum* isolates in Danish small grain cereals. Int J Food Microbiol **157:**384–392. http://dx.doi.org/10.1016/j.ijfoodmicro.2012.06.010.
39. **Charlton ND, Craven KD, Mittal S, Hopkins AA, Young CA.** 2012. *Epichloë canadensis*, a new interspecific epichloid hybrid symbiotic with Canada wildrye (*Elymus canadensis*). Mycologia **104:**1187–1199. http://dx.doi.org/10.3852/11-403.
40. **Pan J, Bhardwaj M, Faulkner JR, Nagabhyru P, Charlton ND, Higashi RM, Miller A-F, Young CA, Grossman RB, Schardl CL.** 2014. Ether bridge formation in loline alkaloid biosynthesis. Phytochemistry **98:**60–68. http://dx.doi.org/10.1016/j.phytochem.2013.11.015.
41. **Takach JE, Young CA.** 2014. Alkaloid genotype diversity of tall fescue endophytes. Crop Sci **54:**667–678. http://dx.doi.org/10.2135/cropsci2013.06.0423.
42. **Schardl CL, Young CA, Moore N, Krom N, Dupont P-Y, Pan J, Florea S, Webb JS, Jaromczyk J, Jaromczyk JW, Cox MP, Farman ML.** 2014. Genomes of plant-associated Clavicipitaceae. Adv Bot Res **70:**291–327. http://dx.doi.org/10.1016/B978-0-12-397940-7.00010-0.
43. **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. http://dx.doi.org/10.1016/S0953-7562(09)80509-1.
44. **Schardl CL, Young CA, Faulkner JR, Florea S, Pan J.** 2012. Chemotypic diversity of epichloae, fungal symbionts of grasses. Fungal Ecol **5:**331–344. http://dx.doi.org/10.1016/j.funeco.2011.04.005.

4 Characterisation of PerA-ΔR proteins

Acknowledgment of assistance:

The results presented in this Chapter 4 were generated with the assistance of Katrin Grage (Massey University) and Wade Mace (AgResearch). The candidate (Daniel Berry) was responsible for all experimental design, bioinformatic analyses and generated all of the pDB plasmids used in this study. The candidate also performed the initial *P. paxilli* transformation, screening and metabolite extraction experiments to validate experimental design. The candidate was responsible for all interpretations of the results presented here. Katrin Grage was responsible for all *P. paxilli* transformations, screening (i.e. all transformations and screening involving plasmids pDB29, pDB30, pDB31, pDB32, pDB34, pDB35 and pDB43) and metabolite extractions following the initial validation experiments. Wade Mace performed all peramine and metabolite analyses on extraction solutions. The assistance of these two co-authors is gratefully acknowledged.

4.1 Hypothesis: PerA-ΔR proteins remain functional

While numerous different mutations resulting in peramine-negative *perA* alleles were identified and described in Chapter 3, the phylogenetic distribution of any given mutation was extremely narrow, and always contained within a single *Epichloë* sp. (Berry et al., 2015b). The only exception was the *perA-ΔR* allele, which was identified in the sexual species *E. festucae*, *E. baconii*, *E. bromicola*, *E. typhina*, *E. typhina* ssp. *clarkii*, *E. typhina* ssp. *poae* and *E. sylvatica*, as well as the asexual species *E. siegelii* and *E.* sp. e4768 (Berry et al., 2015b). This widespread distribution indicated that the *perA-ΔR* allele was inherited by these *Epichloë* spp. from a common ancestor, and the high conservation of the *perA-ΔR* coding sequences (excluding the deleted region encoding the R-domain) suggested these alleles might encode functional proteins (Berry et al., 2015b). It was shown that regions within these *perA-ΔR* alleles were phylogenetically more closely related to all other *perA-ΔR* alleles than to normal *perA* alleles from the same species in an apparent example of trans-species polymorphism (TSP; Berry et al., 2015b). The maintenance of TSPs for extended periods after speciation events requires balancing selection to maintain the divergent alleles within both populations (Tesicky and Vinkler, 2015), and such selection would not occur for *perA-ΔR* alleles if they encoded non-functional proteins. Additionally, RNA-seq data generated from *E. festucae* E2368-infected *Festuca pratensis* (meadow fescue) plants indicated that levels of *perA-ΔR* transcription remained comparable to expression of other bioprotective SM genes (Chris Schardl, personal communication). In the absence of a R-domain PerA-ΔR proteins were not expected to make peramine, and were instead hypothesised to produce a novel dipeptide.

4.2 Results: Bioinformatic analysis of PerA-ΔR proteins

4.2.1 PerA-ΔR proteins contain unique changes to their A1-domain substrate-binding residues

Alignment of protein sequences encoded by *perA* and *perA-ΔR* alleles from a variety of different *Epichloë* spp. identified PerA-ΔR-unique sequence changes located between the 4th and 5th conserved motifs (Marahiel et al., 1997) of the first

AMPylation (A1) domain (Figure 4.1). This region contains the residues responsible for determining A-domain substrate specificity (Stachelhaus et al., 1999). Many of the PerA-ΔR-unique substitutions affected the identity of predicted substrate-binding residues (Challis et al., 2000; Roettig et al., 2011), resulting in alteration of the A1-domain 10AA substrate-specificity code of PerA-ΔR proteins relative to PerA (Figure 4.1). PerA-ΔR proteins from *E. baconii* isolates did not contain these changes. These observations suggested that the A1-domain substrate specificity of PerA-ΔR proteins might be altered relative to PerA.

4.2.2 PerA-ΔR proteins contain additional unique changes to the A2/M-domain junction and T2-domain

Representative PerA and PerA-ΔR sequences were computationally compared to identify regions containing primary sequence changes unique to all PerA-ΔR proteins beyond the A1-domain substrate-binding region. An alignment of the *E. festucae* Fl1 PerA and *E. festucae* E2368 PerA-ΔR protein sequences was also computationally analysed to determine the % protein identity within a 100 aa-wide window in sequential 50 aa steps across the entire PerA/PerA-ΔR alignment. The results showed that PerA-ΔR-unique residues were restricted to the A1-domain substrate-binding region, methylation (M)-domain and second thiolation (T2)-domain (Figure 4.2). These unique residues were located within regions of hyper-variability between the protein sequences of *E. festucae* Fl1 PerA and *E. festucae* E2368 PerA-ΔR (Figure 4.2). An additional hyper-variable region not associated with PerA-ΔR-unique residues was also identified within the condensation (C)-domain (Figure 4.2). This pattern was not observed when the Fl1 PerA sequence was compared to *E. amarillans* E57 PerA (9.4.1), or PerA from the more distantly related *E. typhina* E8 PerA (9.4.2). Figures 4.3 and 4.4 show the location and identity of these PerA-ΔR-unique changes on an alignment of representative PerA and PerA-ΔR protein sequences for the M and T2-domains, respectively.

4.2.3 Evidence for recombination between *perA* and *perA-ΔR* alleles

The comparison of *perA*/*perA-ΔR* alleles presented in Berry et al. (2015b) revealed that the phylogenetic topology of T2-domain-encoding sequences was

consistent with TSP, whereas the topology of the A2-domain-encoding sequences roughly matched species phylogeny. This disparity was hypothesised to be caused by recombination between ancestral *perA* and *perA-ΔR* alleles. Representative *perA* and *perA-ΔR* alelles from *E. festucae*, *E. baconii*, *E. bromicola*, *E. typhina* and *E. typhina* ssp. *poae* were analysed using the DualBrothers software package (Minin et al., 2005) for evidence of recombination. The results showed that recombination events were predicted to have occurred between *perA* and *perA-ΔR* alleles in both *E. festucae* and *E. bromicola* around the A1-domain substrate-binding region, the N-terminus of the M-domain and the entire T2-domain, resulting in the *perA-ΔR* alleles observed in these isolates today (Figure 4.5). A similar pattern of recombination between *perA* and *perA-ΔR* alleles was observed in *E. baconii perA-ΔR* alleles, with the exception that no recombination event was observed for the A1-domain substrate-binding region (Figure 4.5). Non-synonymous SNPs were also computationally removed from an alignment of representative *E. festucae* and *E. typhina perA*/*perA-ΔR* alleles to attempt to determine if evolutionary convergence in *perA-ΔR* alleles could be the cause of these predicted recombination events. The results showed that with non-synonymous SNPs removed, the calculated probability for the *perA-ΔR*-grouped topology dropped considerably, but was not eliminated (Figure 4.5D). This low signal was due to the low number of synonymous SNPs present at these locations. However, given that *perA-ΔR*-unique synonymous SNPs were not observed at any other locations across these alignments, these results are still proposed to support the hypothesis that *perA-ΔR* alleles in *E. festucae*, *E. baconii* and *E. bromicola* are the result of recombination between ancestral *perA* and *perA-ΔR* alleles.

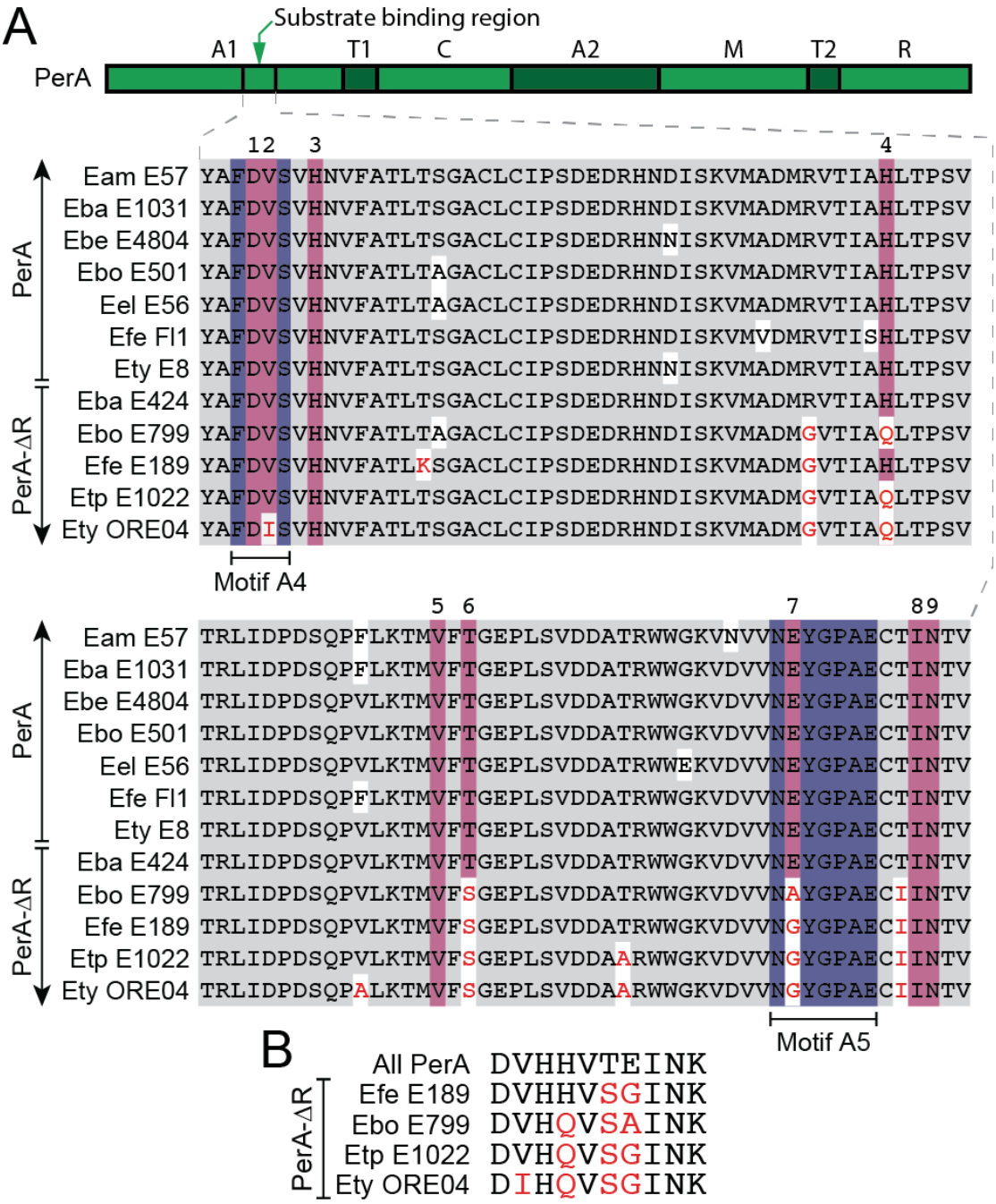


Figure 4.1 Comparison of the A1-domain substrate-binding region between PerA and PerA-ΔR proteins. (A) Multiple sequence alignment of the first AMPylation (A1)-domain substrate-binding region from representative PerA and PerA-ΔR sequences. Alignment location is indicated on the PerA protein sequence map shown at top. PerA-ΔR-unique residues are coloured red. Conserved A-domain A4/A5 motifs (blue highlights) and substrate-binding residues (purple highlights) are based on findings of previous studies (Conti et al., 1997; Lee et al., 2010; Marahiel et al., 1997; Stachelhaus and Marahiel, 1995; Stachelhaus et al., 1999). Abbreviations are as follows: Eam, *E. amarillans*; Eba, *E. baconii*; Ebe, *E. brachyelytri*; Ebo, *E. bromicola*; Eel, *E. elymi*; Efe, *E. festucae*; Ety, *E. typhina*; Etp, *E. typhina* ssp. *poae*. (B) 10AA substrate specificity code of PerA and PerA-ΔR proteins.

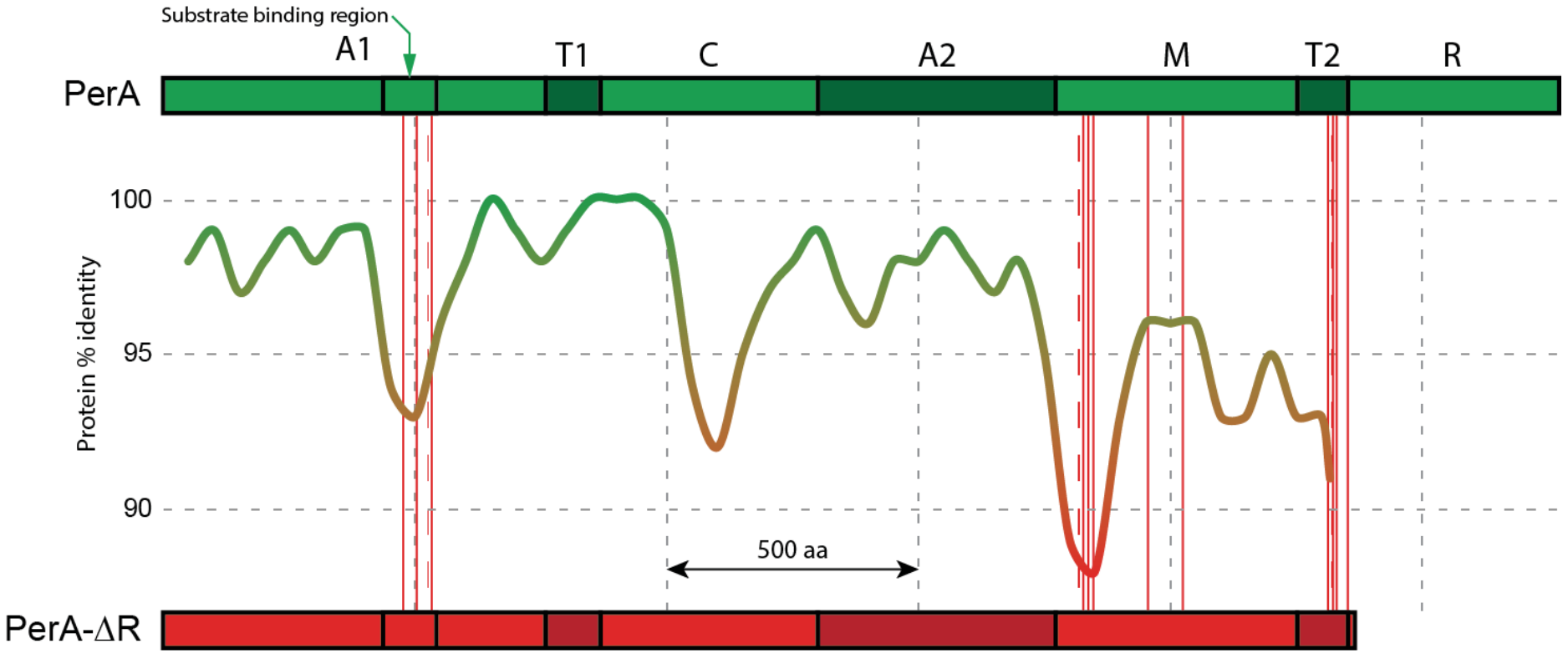


Figure 4.2 Comparison of PerA and PerA-ΔR protein sequences. Protein sequence maps of PerA and PerA-ΔR are separated by a graph showing the % identity from a 100 aa window sliding analysis (50 aa step size) of the *E. festucae* Fl1 PerA and *E. festucae* E2368 PerA-ΔR protein sequences. Vertical red bars show the positions of residues identified as PerA-ΔR-unique when PerA-ΔR protein sequences from *E. festucae*, *E. bromicola*, *E. typhina* and *E. typhina* ssp. *poae* were compared to representative peramine-producing PerA sequences. Solid bars indicate positions where the residue identity was the same across all PerA-ΔR sequences analysed, while dashed bars indicate positions where the residue identity differed between the PerA-ΔR sequences analysed, but still remained different to all PerA sequences analysed. The PerA protein sequence map is annotated with approximate domain boundaries with abbrieviations as described previously (Section 1.2).

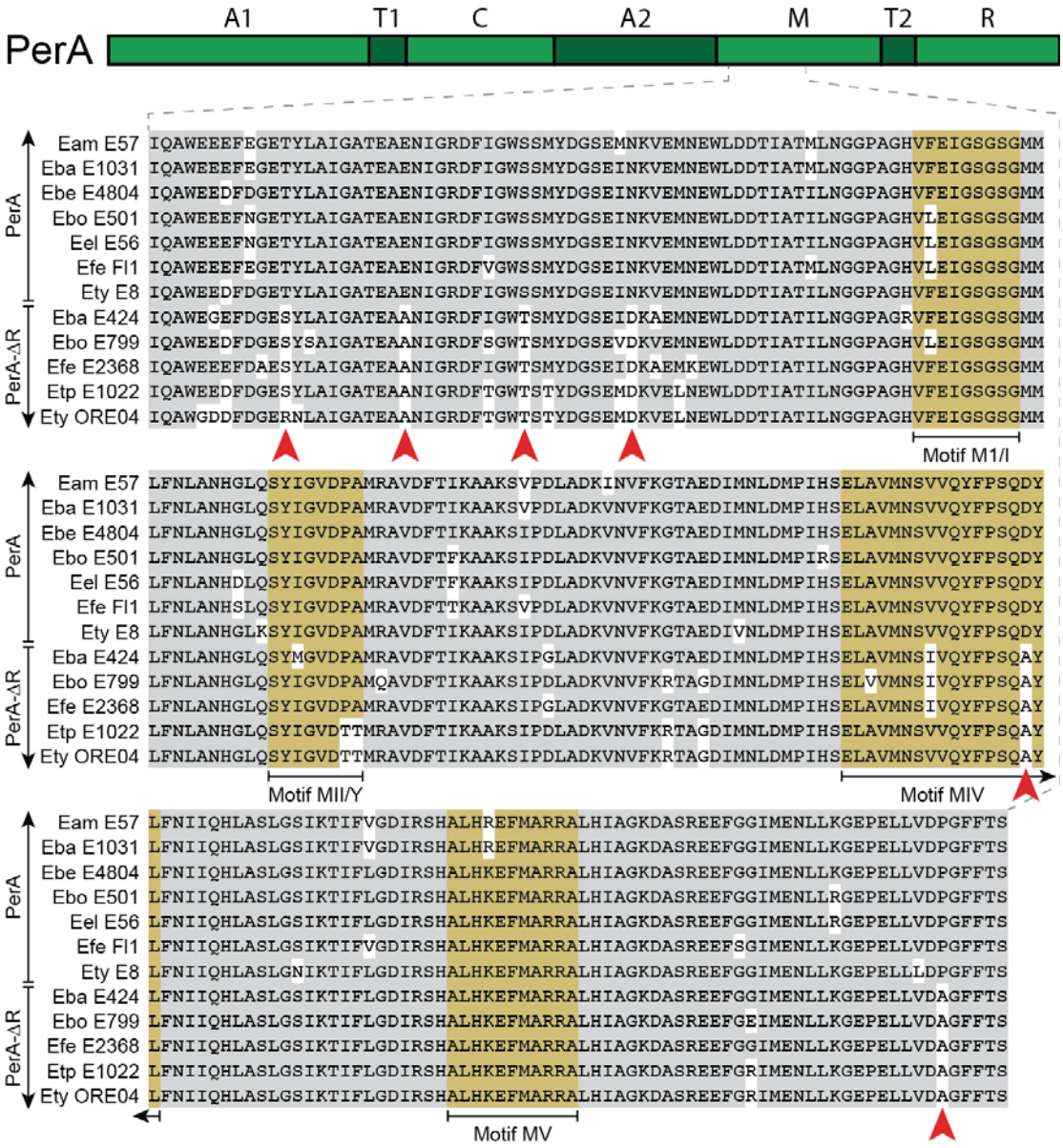


Figure 4.3 Multiple sequence alignment of M-domain. Alignment of PerA/PerA-ΔR sequences from representative isolates showing the N-terminal half of the M-domain region, with PerA-ΔR-unique substitutions annotated with red arrowheads. *N*-methyltransferase conserved motifs reported in previous studies are highlighted in yellow (Ansari et al., 2008; Hacker et al., 2000).

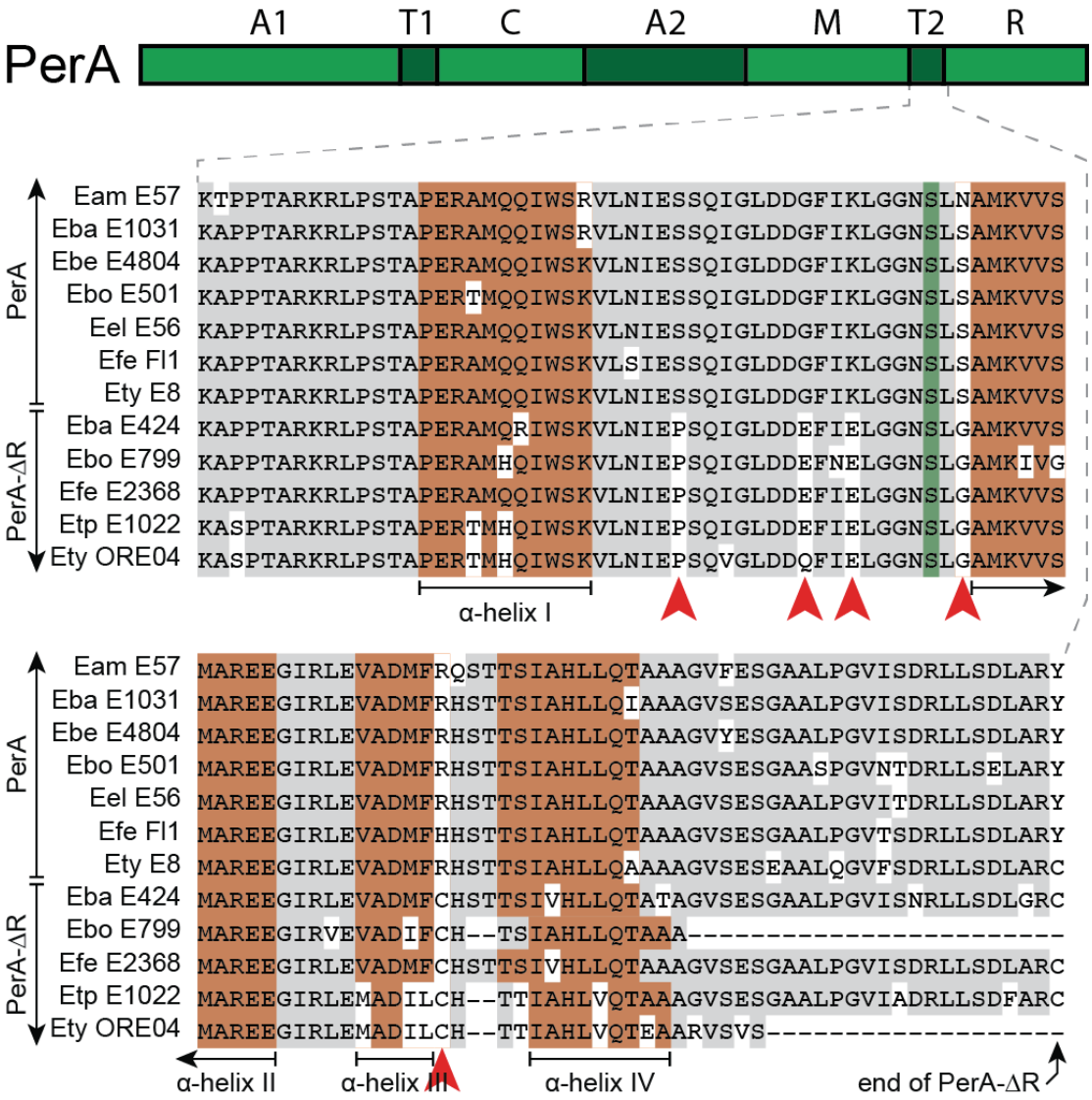


Figure 4.4 Multiple sequence alignment of T2-domain. Alignment of PerA/PerA-ΔR T2-domain sequences from representative isolates. PerA-ΔR-unique residues are annotated with red arrowheads. Orange boxes show the locations of α-helices based on structural predictions (Drozdetskiy et al., 2015; Kelley et al., 2015) and comparison to previously solved T-domains (Haslinger et al., 2015; Lohman et al., 2014; Tufar et al., 2014). The serine residue predicted to be targeted for post-translational modification by a 4'-phosphopantetheinyl transferase is highlighted in green.

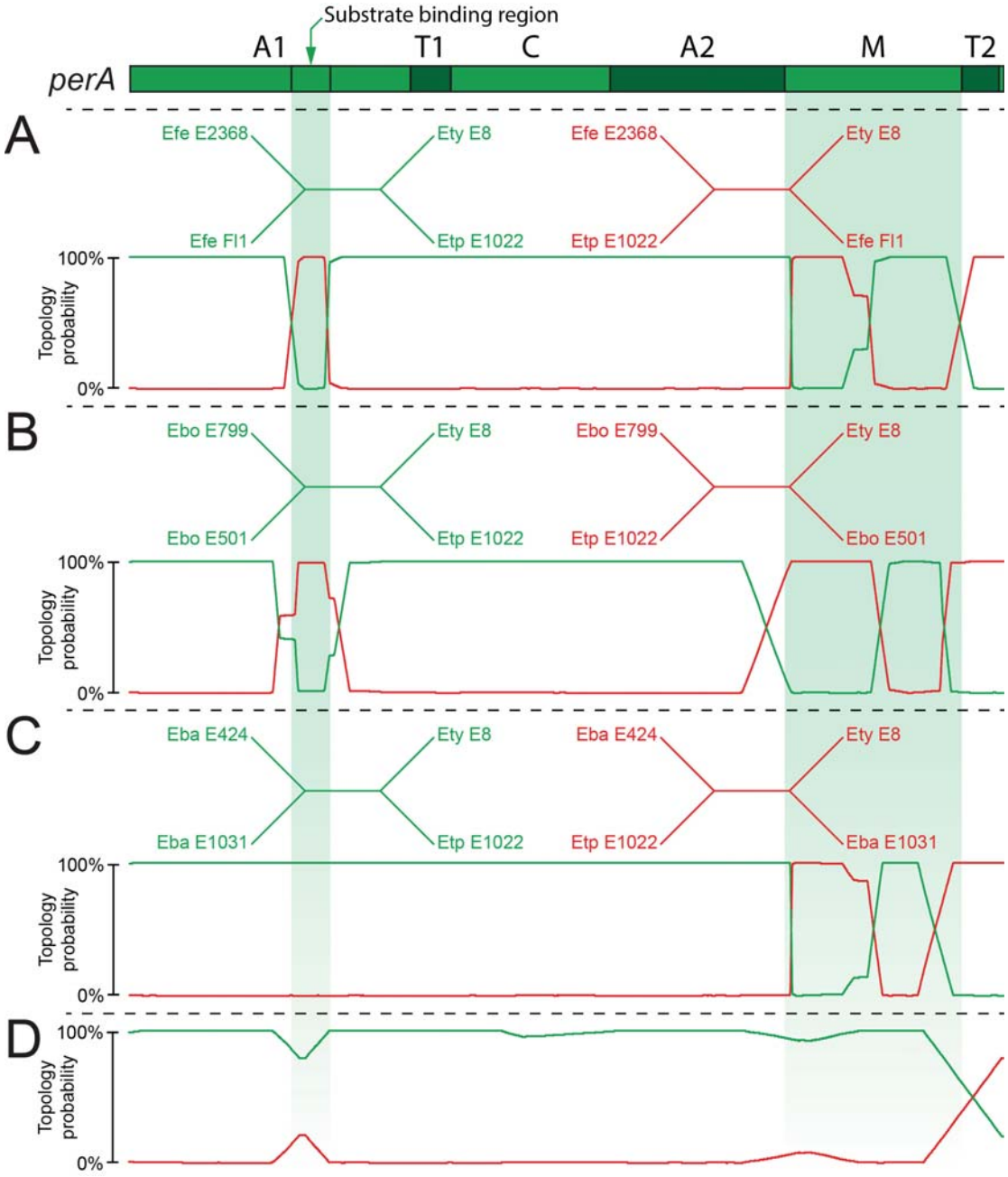


Figure 4.5 Evidence for recombination between *perA* and *perA-ΔR* alleles. Graph traces show the calculated probability for a particular phylogenetic topology, as described by the equivalently coloured phylograms displayed above each graph, in a sliding window analysis and are plotted to scale relative to the truncated *perA* gene map shown at the top of the figure. *E. typhina* E8 *perA* and *E. typhina* ssp. *poae* E1022 *perA-ΔR* sequences were compared to (A) *E. festucae* Fl1 *perA*/*E. festucae* E2368 *perA-ΔR*, (B) *E. bromicola* E501 *perA*/*E. bromicola* E799 *perA-ΔR* or (C) *E. baconii* E1031 *perA*/*E. baconii* E424 *perA-ΔR*. Comparisons used the DualBrothers recombination detection software (Minin et al., 2005) to infer changes in topology (using settings: chain length 2×10^6, subsampling frequency 10, burn-in length 1×10^4, window length 100, step size 10 using all possible topologies setting). (D) Predicted topology of the same sequences as (A), but with all non-synonymous SNPs computationally removed (phylograms for D are the same as A).

4.2.4 Analysis of the PerA A1-domain N-terminal extension

The PerA and PerA-ΔR proteins contain a 250 aa region upstream of the A1-domain for which the function was unknown. A recently published study by Kalb et al. (2015) analysed a similar A-domain N-terminal extension in the NRPS-like adenylate-forming reductase protein NPS3 from the basidiomycete *Ceriporiopsis subvermispora*, and found that this N-terminal extension was required for substrate AMPylation by the NPS3 A-domain. The sequence of this NPS3 N-terminal extension was shown to align well with the second subunit of a NRPS C-domain (Kalb et al., 2015). The 250 aa N-terminal extension of PerA was therefore analysed using the Phyre2 protein fold recognition server (Kelley et al., 2015), and was also predicted to be structurally closely related to subunit 2 of a NRPS C-domain.

4.3 Results: Establishment of *Penicillium paxilli* as a heterologous expression host for *perA* genes

4.3.1 Rationale for heterologous gene expression in *P. paxilli*

The amplitude of *Epichloë* bio-protective SM gene expression is generally many orders of magnitude higher *in planta* than in culture (Chujo and Scott, 2014; Eaton et al., 2010; Eaton et al., 2015; Fleetwood et al., 2007; Spiering et al., 2005; Young et al., 2005; Young et al., 2006), and *perA* is no exception to this trend (Tanaka et al., 2005). Analysing the products produced by modified *perA* or *perA-ΔR* alleles expressed in *E. festucae* would therefore have been very time consuming, as transformants would first have to be inoculated into plants. Additionally, the plant metabolite background would have complicated identification of unknown compounds. An alternative would have been to place the altered *perA*/*perA-ΔR* alleles under the control of a promoter that initiates strong, constitutive expression to allow expression in culture. However, *E. festucae* is a relatively slow-growing fungus, and purification of *E. festucae* transformants is complicated by a lack of sporulation in culture. It was also hypothesised that strong expression of a secondary metabolite gene well before stationary phase might adversely affect culture growth. It was therefore decided to express these genes in *Penicillium paxilli*, which had previously been shown to be a suitable heterologous host for *E. festucae*

SM genes (Young et al., 2006). The genome of *P. paxilli* was therefore sequenced to confirm it contained intact genes encoding a 4'phosphopantetheinyl transferase and a type-II thioesterase, as both are required for NRPS function.

4.3.2 Draft genome sequence of the filamentous fungus *Penicillium paxilli* (ATCC 26601)

MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
TO DOCTORAL THESIS CONTAINING PUBLICATIONS

(To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis)

We, the candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: Daniel Berry

Name/Title of Principal Supervisor: Barry Scott

Name of Published Research Output and full reference:

Berry, D., Cox, M.P., and Scott, B. (2015). Draft genome sequence of the filamentous fungus Penicillium paxilli (ATCC 26601). Genome Announcements 3, e00071-00015.

In which Chapter is the Published Work: 4

Please indicate either:

- The percentage of the Published Work that was contributed by the candidate: 50%

  and / or

- Describe the contribution that the candidate has made to the Published Work:

  Generation of genomic DNA for sequencing, all analysis of the resulting genome post-assembly, and writing of the paper.

Daniel Berry — Digitally signed by Daniel Berry DN: cn=Daniel Berry, o=Massey University, ou=Institute of Fundamental Sciences, email=D.Berry@massey.ac.nz, c=NZ Date: 2016.03.27 19:40:02 +13'00'

Candidate's Signature — 27/3/16 — Date

Barry Scott — Digitally signed by Barry Scott DN: cn=Barry Scott, o=Massey University, ou=IFS, email=d.b.scott@massey.ac.nz, c=NZ Date: 2016.03.27 17:30:37 +02'00'

Principal Supervisor's signature — 27/3/16 — Date

Draft Genome Sequence of the Filamentous Fungus *Penicillium paxilli* (ATCC 26601)

Daniel Berry, Murray P. Cox, Barry Scott

Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand

***Penicillium paxilli* ATCC 26601 is an asexual filamentous fungal species known for its production of the mycotoxin paxilline. We present here the 35-Mb draft genome sequence for this organism.**

Received 19 January 2015 **Accepted** 4 February 2015 **Published** 12 March 2015

Citation Berry D, Cox MP, Scott B. 2015. Draft genome sequence of the filamentous fungus *Penicillium paxilli* (ATCC 26601). Genome Announc 3(2):e00071-15. doi:10.1128/genomeA.00071-15.

Address correspondence to Barry Scott, d.b.scott@massey.ac.nz.

Penicillium paxilli is an asexual saprophytic fungus, and isolate ATCC 26601 originates from insect-damaged pecans in Georgia, USA (1). *P. paxilli* is known for its production of the indole-diterpene paxilline, a potent tremorgen (1) that inhibits calcium-activated potassium channels (2). Due to the high rate of growth, sporulation, and ease of genetic modification, *P. paxilli* has been utilized as a model organism for studying indole-diterpene biosynthesis (3–8). The paxilline biosynthetic machinery is encoded by the *PAX* gene cluster, and the promoter from the *paxM* gene of this cluster has been successfully used for the heterologous expression of indole-diterpene biosynthetic genes from the fungal grass endophyte *Epichloë festucae* (9). The genome sequence of *P. paxilli* was sequenced to investigate its suitability as a heterologous expression system for fungal nonribosomal peptide synthetases (NRPS).

P. paxilli genomic DNA was prepared by phenol-chloroform extraction (10) and treated with RNase A. Two runs with 150-bp paired-end fragment reads were done on an Illumina MiSeq instrument by New Zealand Genomics Limited (NZGL), obtaining approximately 182-fold genome coverage. The reads were dynamically trimmed using the SolexaQA package to their longest fragment, such that the base call error rates did not exceed a *P* value of 0.05, and paired-end reads of <100 bp were discarded. *De novo* assembly was performed on all data using ABySS version 1.3.0, with the parameters $n = 2$, $c = 10$, and $k = 79$. The final assembly consisted of 635 contigs, with 414 contigs >500 bp in size, an average contig length of 84,079 bp, a maximum contig length of 732,567 bp, an N_{50} of 189,821 bp, and a total of 34,808,516 residues.

Blast+ was used to identify a 4′-phosphopantetheinyl transferase (DDBJ/EMBL/GenBank accession no. KP233470) and a type II thioesterase (accession no. KP233471) in the *P. paxilli* genome that are likely to be functional. These proteins are required to produce and maintain functional NRPS enzymes, indicating that *P. paxilli* has the potential to be a suitable heterologous expression system for fungal NRPSs.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. AOTG00000000. The version described in this paper is the first version, AOTG01000000.

ACKNOWLEDGMENTS

This research was supported by a Massey University Ph.D. scholarship and a grant from the Royal Society of New Zealand Marsden Fund (contract MAU1002).

REFERENCES

1. **Cole RJ, Kirksey JW, Wells JM.** 1974. A new tremorgenic metabolite from *Penicillium paxilli*. Can J Microbiol **20:**1159–1162. http://dx.doi.org/10.1139/m74-179.
2. **Knaus HG, McManus OB, Lee SH, Schmalhofer WA, Garcia-Calvo M, Helms LM, Sanchez M, Giangiacomo K, Reuben JP, Smith AB, Kaczorowski GJ, Garcia ML.** 1994. Tremorgenic indole alkaloids potently inhibit smooth muscle high-conductance calcium-activated potassium channels. Biochemistry **33:**5819–5828. http://dx.doi.org/10.1021/bi00185a021.
3. **Saikia S, Nicholson MJ, Young C, Parker EJ, Scott B.** 2008. The genetic basis for indole-diterpene chemical diversity in filamentous fungi. Mycol Res **112:**184–199. http://dx.doi.org/10.1016/j.mycres.2007.06.015.
4. **Saikia S, Parker EJ, Koulman A, Scott B.** 2006. Four gene products are required for the fungal synthesis of the indole-diterpene, paspaline. FEBS Lett **580:**1625–1630. http://dx.doi.org/10.1016/j.febslet.2006.02.008.
5. **Saikia S, Parker EJ, Koulman A, Scott B.** 2007. Defining paxilline biosynthesis in *Penicillium paxilli*: functional characterization of two cytochrome P450 monooxygenases. J Biol Chem **282:**16829–16837. http://dx.doi.org/10.1074/jbc.M701626200.
6. **Saikia S, Scott B.** 2009. Functional analysis and subcellular localization of two geranylgeranyl diphosphate synthases from *Penicillium paxilli*. Mol Genet Genomics **282:**257–271. http://dx.doi.org/10.1007/s00438-009-0463-5.
7. **Scott B, Young CA, Saikia S, McMillan LK, Monahan BJ, Koulman A, Astin J, Eaton CJ, Bryant A, Wrenn RE, Finch SC, Tapper BA, Parker EJ, Jameson GB.** 2013. Deletion and gene expression analyses define the paxilline biosynthetic gene cluster in *Penicillium paxilli*. Toxins **5:**1422–1446. http://dx.doi.org/10.3390/toxins5081422.
8. **Young C, 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. http://dx.doi.org/10.1046/j.1365-2958.2001.02265.x.
9. **Young CA, Felitti S, Shields K, Spangenberg G, Johnson RD, Bryan GT, Saikia S, Scott B.** 2006. A complex gene cluster for indole-diterpene biosynthesis in the grass endophyte *Neotyphodium lolii*. Fungal Genet Biol **43:**679–693. http://dx.doi.org/10.1016/j.fgb.2006.04.004.
10. **Byrd AD, Schardl CL, Songlin PJ, Mogen KL, Siegel MR.** 1990. The β-tubulin gene of *Epichloë typhina* from perennial ryegrass (*Lolium perenne*). Curr Genet **18:**347–354. http://dx.doi.org/10.1007/BF00318216.

4.3.3 Expression of *perA* in *P. paxilli* as a pilot study

To investigate whether *P. paxilli* was a suitable host for expressing *perA* and *perA-ΔR* alleles the *E. festucae* Fl1 *perA* gene, which encodes a PerA protein known to produce peramine (Tanaka et al., 2005), was inserted by yeast recombinational cloning (2.6.11) between the promoter and terminator sequences of the *P. paxilli* SM gene *paxM* to generate plasmid pDB05 (9.3.6). Sequence fidelity of the insert was confirmed by sequencing (2.6.13). Previous studies had shown that P*paxM* was effective at promoting expression of *E. festucae* genes in *P. paxilli* (Saikia et al., 2012; Young et al., 2006). Additionally, the involvement of the protein encoded by *paxM* in biosynthesis of the secondary metabolite paxilline (Saikia et al., 2006) suggested that P*paxM*-driven expression would be initiated only at late exponential phase, making this promoter ideal for expression of a SM gene like *perA* after biomass accumulation. Plasmid pDB05 was co-transformed into *P. paxilli* protoplasts alongside pAN7-1 (9.3.3), which provided a selectable marker for resistance to the antibiotic hygromycin. Crude gDNA extracts (2.5.6) from these transformants were then used as template for PCR amplification (2.6.3) with primer pairs p9f/p9r, PPperA_SCR_1/PPperA_SCR_2 and PPperA_SCR_3/PPperA_SCR_4, which amplify within the A2-domain encoding region of *perA*, from within *perA* across P*paxM* and from within *perA* across T*paxM*, respectively (9.3.6). Two transformants (strains 05-31 and 05-33) from which extracted gDNA was successfully used as template for PCR amplification using each of the above primer pairs were then grown in liquid cultures (2.4.3.1) alongside wt for 72 h. The mycelia harvested (2.4.4) from these cultures were then used for RNA extraction (2.7.1). These RNA samples were used for RT-PCR analysis (2.7.3; 2.6.3) with primer pair p9f/p9r, and the results confirmed that *perA* was expressed in both transformants (Figure 4.6). Because the *perA* gene does not contain any introns RT-PCR using primers tub2_F/tub2_R, which bind within exons of the *P. paxilli* β-tubulin gene enabling PCR amplification across two introns to produce a noticeably larger product from gDNA vs cDNA template, was performed on each RNA sample to check for gDNA contamination (Figure 4.6). Duplicate liquid cultures (2.4.3.1) of strains 05-31 and 05-33 were grown alongside wt *P. paxilli* for 168 h at 22°C/200 RPM. A second set of duplicate cultures were grown under the same conditions, except that the CDYE media was supplemented with 0.5 g L^{-1}

proline and 0.5 g L^{-1} arginine to determine if this would increase peramine production. Arginine and 1-pyrroline-5-carboxylate (P5C) are the predicted PerA substrates (Tanaka et al., 2005). Proline was used instead of P5C as the latter could not be sourced; P5C is an intermediate in the bidirectional arginine and proline metabolic pathway (Aral and Kamoun, 1997), though the actual effect of increased proline concentration on the levels of P5C available was not tested. Mycelia from these cultures were isolated (2.4.4), freeze-dried, and peramine was extracted from 20 mg of lyophilised mycelia from each sample (2.11.1). The peramine content of these extracts was then measured by LCMS (2.11.2). The results showed that peramine was not detected in mycelia from wt *P. paxilli*, while strains 05-31 and 05-33 produced approx. 3-5 ppm peramine (Table 4.1). No obvious difference was observed between cultures grown in CDYE media with or without supplemental Pro/Arg (Table 4.1). After peramine production was confirmed for strains 05-31 and 05-33, a third pDB05 transformant (05-43) was identified using the screening method described above to allow the use of biological triplicates in future experiments (Figure 4.6).

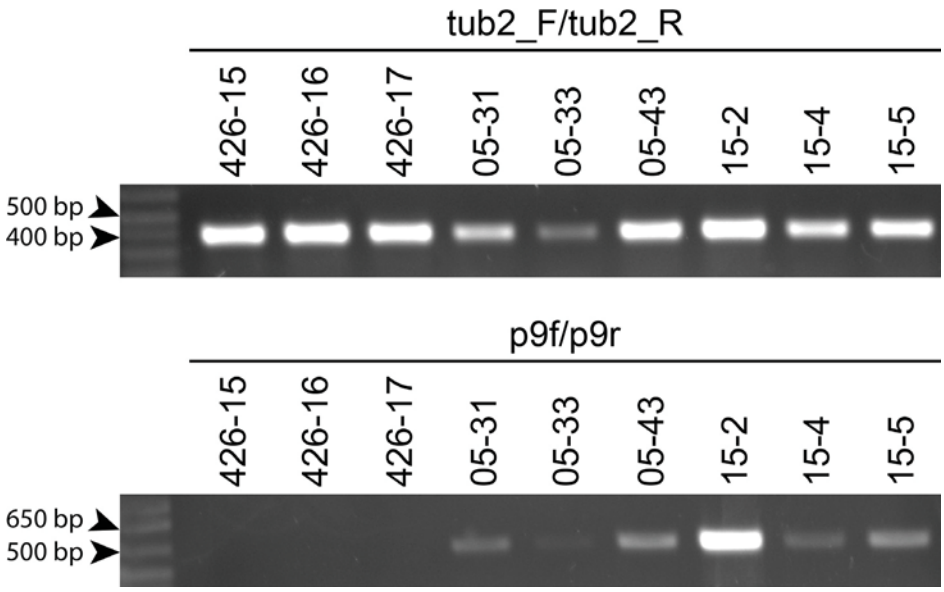


Figure 4.6 RT-PCR analysis of *P. paxilli* pDB05, pDB15 and pRS426 transformants. Gels show amplification of 1:5-diluted cDNA using primer pair tub2_F/tub2_R, which amplifies within the *P. paxilli* β-tubulin gene (expected product size 532 bp for gDNA template and 395 bp for cDNA template), or p9f/p9r (expected product size 606 bp), which amplifies within the A2-domain encoding region of *perA*. 1 kb plus ladder (Invitrogen) shown in first lane of both gels with relevant size indicators annotated.

Table 4.1 Peramine production in *P. paxilli perA* transformants

Strain	Culture #	Media	[Peramine] (ppm)
wt	1	CDYE	0
wt	2	CDYE	0
wt	1	CDYE + 0.5 g L^{-1} Pro, Arg	0
wt	2	CDYE + 0.5 g L^{-1} Pro, Arg	0
05-31	1	CDYE	3.8
05-31	2	CDYE	2.9
05-31	1	CDYE + 0.5 g L^{-1} Pro, Arg	4.2
05-31	2	CDYE + 0.5 g L^{-1} Pro, Arg	3.7
05-33	1	CDYE	4.6
05-33	2	CDYE	4.9
05-33	1	CDYE + 0.5 g L^{-1} Pro, Arg	4.5
05-33	2	CDYE + 0.5 g L^{-1} Pro, Arg	5.1

4.3.4 Optimisation of culture growth time for peramine production in *P. paxilli*

The growth period used for of *P. paxilli* liquid cultures in these initial experiments (168 h/7 d) was based on methods from previous paxilline studies (Saikia et al., 2006; Young et al., 2006). Paxilline production requires the protein encoded by *paxM* for biosynthesis (Saikia et al., 2006), suggesting that similar growth conditions would be suitable for production of peramine in *P. paxilli* with a P*paxM-perA* expression construct. It had been previously shown that while the dry mass of *P. paxilli* mycelia grown in CDYE liquid media decreases after peaking at approx. 48 h growth time, paxilline was not detectable until the cultures had been growing for 72 h (Telfer, 2000). Paxilline then continued to accumulate until at least 144 h, at which point paxilline concentrations had increased approx. 20-fold over the 72 h levels (Telfer, 2000). Paxilline was also detected in the media only after 144 h, indicating that significant cell autolysis may begin around this time point (Telfer, 2000). A time-course analysis of peramine production was therefore performed over 168 h using the *perA*-expressing strain 05-33 to determine the optimal growth time for peramine production. Both freeze-dried mycelia and filtered media samples were used for this analysis.

Replicate liquid cultures (18 total; 2.4.3.1) of *P. paxilli* strain 05-33 were grown at 22°C/200 RPM. Three cultures were removed from the incubator after 48 h for harvesting of mycelia/media samples (2.4.4), with this process repeated every 24 h thereafter. Peramine was extracted from 20 mg of lyophilised mycelia and 500 µL of media from each sample (2.11.1), and the peramine concentration of each sample was measured (2.11.2). The results showed that, like paxilline, peramine was first detected in the mycelia after 72 h growth (Figure 4.7). Peramine concentrations then increased approx. 5-fold over the next 24 h. However, unlike paxilline, peramine concentrations plateaued after 96 h (Figure 4.7). Peramine was only detected in the media samples after 96 h, with the peramine concentration increasing in the media over the next 48 h and peaking at 144 h. Based on these results, all subsequent *P. paxilli* liquid cultures were grown for 120 h (5 d) to minimise any complications associated with the increased cell autolysis expected between 120-144 h (Telfer, 2000) while still allowing 24 h leeway beyond the 96 hour peramine peak in case of delayed induction.

4.3.5 Optimisation of culture growth temperature for peramine production in *P. paxilli*

Liquid cultures of *P. paxilli* strains for use in initial experiments were incubated at the optimal growth temperature for *E. festucae* (22°C); however, *P. paxilli* cultures for secondary metabolite analyses in previous studies had been incubated at 28°C (Saikia et al., 2006; Young et al., 2006). To determine if temperature significantly impacted peramine production *P. paxilli* strain 05-33 liquid cultures (2.4.3.1) were grown in triplicate for 120 h with shaking of 200 RPM at either 22°C or 28°C. Mycelia were then harvested (2.4.4), freeze-dried and peramine was extracted from 40 mg of lyophlised mycelia from each sample (2.11.1). The peramine content of these extracts was then measured by LCMS (2.11.2). The results showed that cultures grown at 28°C produced significantly less peramine than those grown at 22°C (Figure 4.8; $P < 0.05$). The 22°C incubation temperature was therefore used for all subsequent liquid cultures of *P. paxilli*.

4.3.6 Comparison of peramine production between *P. paxilli* wt and Δ*PAX* strains expressing *perA*

Previous studies have suggested that *P. paxilli* redirects significant resources towards paxilline production during stationary phase (Telfer, 2000). This might limit substrate availability for PerA, resulting in the observed peramine concentration plateau from 96 h onwards. *P. paxilli* strain PN2253 contains a 10-gene deletion within the *PAX* gene cluster (Young et al., 2001), including the four gene products required for synthesis of the first committed pathway intermediate, paspaline (Saikia et al., 2006). To determine if eliminating the paxilline biosynthetic pathway would increase peramine production the *perA* expression plasmid pDB05 (9.3.6) was co-transformed with pII99 (9.3.4) into PN2253 protoplasts (2.9.2). Plasmid pII99 provided a selectable marker for resistance to the antibiotic geneticin, which was necessary as strain PN2253 was already hygromycin resistant. Crude gDNA extracted from these transformants (2.5.6) and used as template for PCR (2.6.3) using primer pairs KG11/KG12, which amplifies the 3' end of *perA*; PPperA_SCR_1/PPperA_SCR_2, which amplifies from within the 5' end of *perA* across the upstream *paxM* promoter sequence; and PPperA_SCR_3/PPperA_SCR_4, which amplifies from within the 3' end *perA* across the downstream *paxM* terminator sequence (9.3.6). Liquid cultures (2.4.3.1) of three transformants that tested positive for all PCR fragments (ΔPAX05-47, ΔPAX05-49 and ΔPAX05-55) were grown for 5 d and RNA was extracted (2.7.1) from mycelia isolated from those cultures (2.4.4). RT-PCR analysis of this RNA (2.7.3; 2.6.3) confirmed expression of *perA* in all three transformants (Figure 4.9). Triplicate liquid cultures (2.4.3.1) of each ΔPAX05 strain were then grown for 120 h at 22°C/200 RPM alongside strain 05-33, mycelia were harvested (2.4.4), freeze-dried and peramine was extracted (2.11.1) from 40 mg of each lyophilised sample. The peramine content of these extracts was then determined by LCMS (2.11.2). The results showed that expression of *perA* in a Δ*PAX* background did not increase peramine production relative to strains expressing *perA* in a wt background, and might even reduce peramine production (Table 4.2). No peramine was detected in mycelia from strain ΔPAX05-55, indicating a *perA* integration problem that was not detected by the PCR or RT-PCR screens.

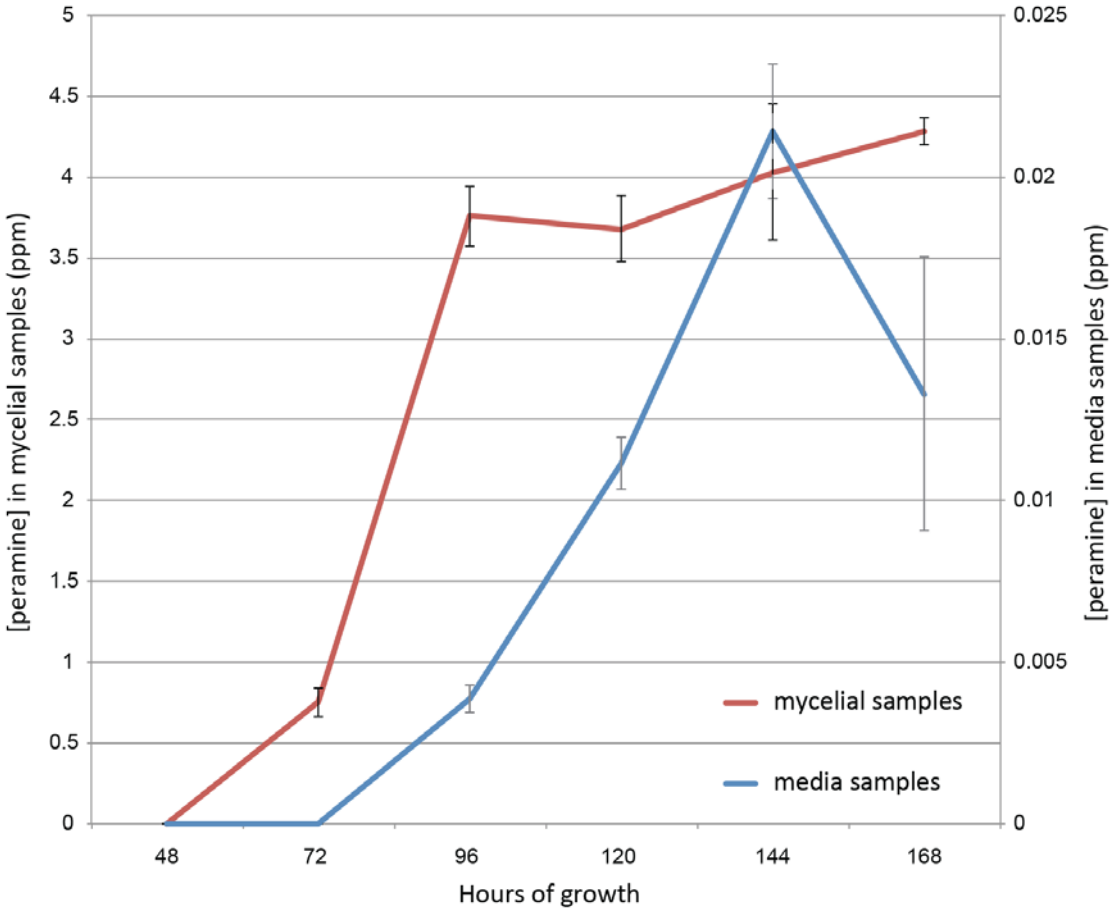


Figure 4.7 **Peramine concentration time course.** Each data point shows the mean of three independent samples, with error bars showing the standard error of each mean. Limit of detection/quantification is 0.1/0.2 ppm for mycelial samples and 0.002/0.006 ppm for media samples.

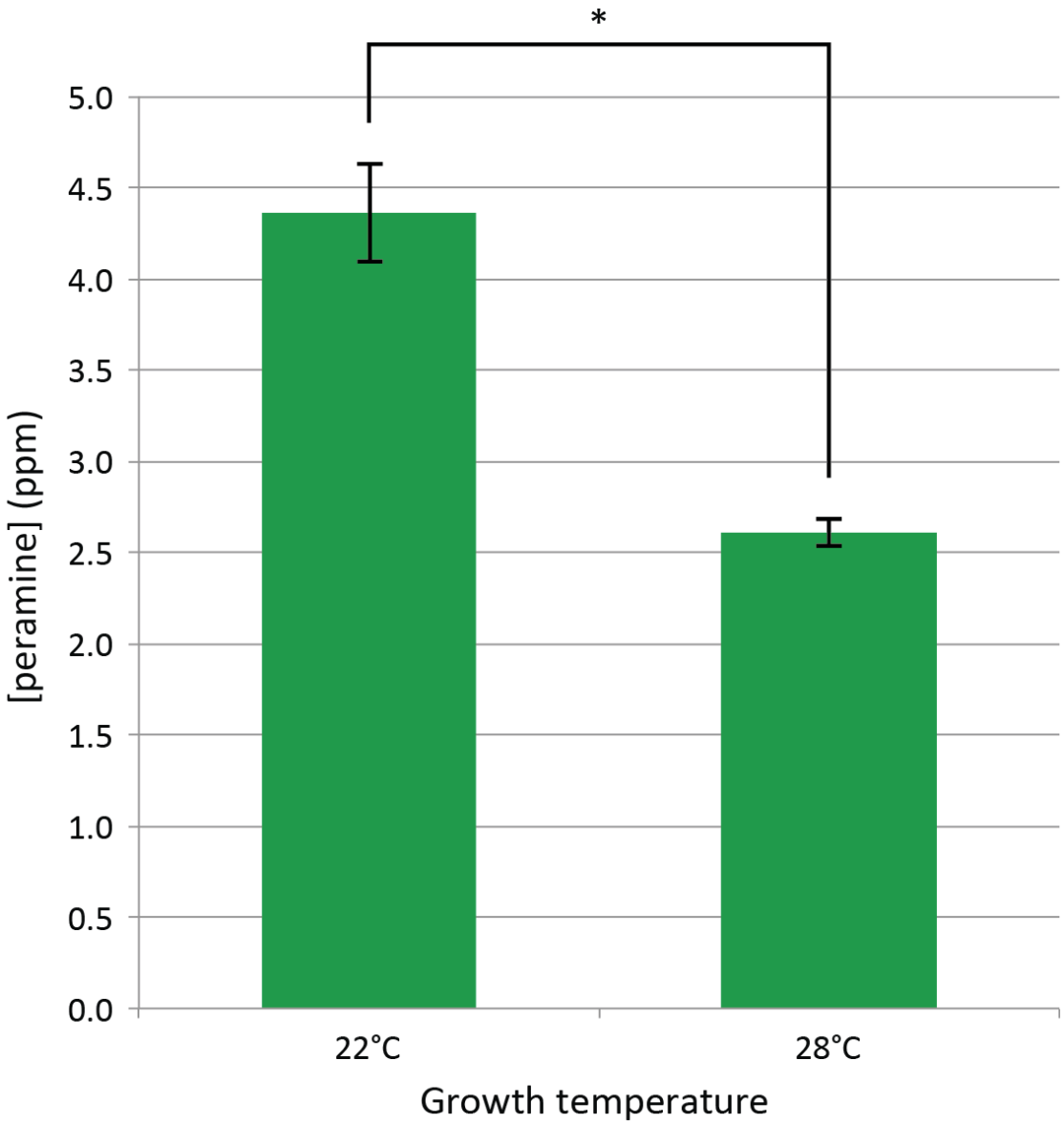


Figure 4.8 Effect of growth temperature on peramine production. Each column shows mean peramine concentration of three independent samples grown at the indicated temperature for 5 d. Error bars show the standard error of the mean. The statistical significance of the difference between these means was calculated using one-tailed Welch's *t*-test (P = 0.013).

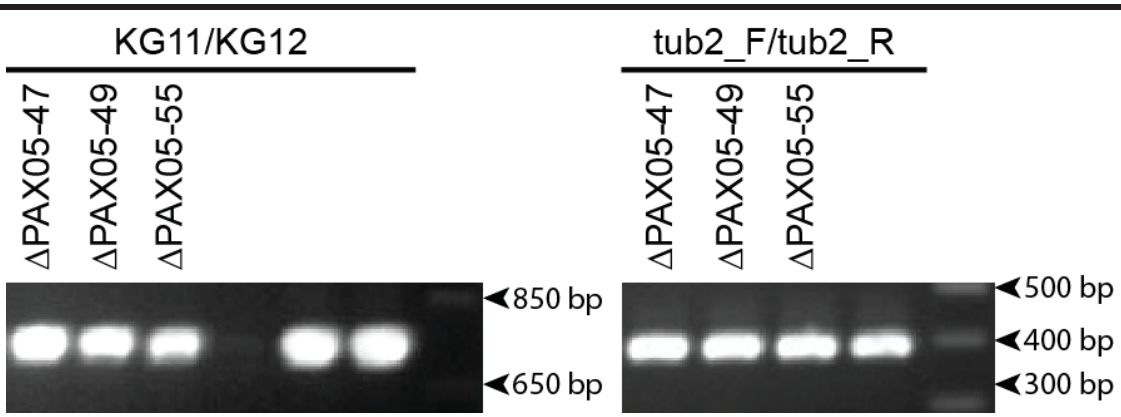


Figure 4.9 RT-PCR analysis of Δ*PAX* pDB05 transformants. Gels show amplification of cDNA using primer pair KG11/KG12 (expected product size 735 bp), which amplifies the 3' end of *perA*, or tub2_F/tub2_R, which amplifies within the *P. paxilli* β-tubulin gene (expected product size 532 bp for gDNA template and 395 bp for cDNA template). 1 kb plus ladder (Invitrogen) shown in last lane of both gels with relevant size indicators annotated.

Table 4.2 Peramine production by ΔPAX05 strains

Strain	Genotype	Culture #	[Peramine] (ppm)
05-33	wt/*perA*	1	4.6
05-33	wt/*perA*	2	3.8
05-33	wt/*perA*	3	4.6
ΔPAX05-47	Δ*PAX*/*perA*	1	1.9
ΔPAX05-47	Δ*PAX*/*perA*	2	1.6
ΔPAX05-47	Δ*PAX*/*perA*	3	0.9
ΔPAX05-49	Δ*PAX*/*perA*	1	3.5
ΔPAX05-49	Δ*PAX*/*perA*	2	3.5
ΔPAX05-49	Δ*PAX*/*perA*	3	2.3
ΔPAX05-55	Δ*PAX*/*perA*	1	0.0
ΔPAX05-55	Δ*PAX*/*perA*	2	0.0
ΔPAX05-55	Δ*PAX*/*perA*	3	0.0

4.4 Results: Functional characterisation of PerA-ΔR proteins

4.4.1 Comparison of A1-domain substrate specificity between PerA and PerA-ΔR proteins

Three constructs containing modified *perA* genes under control of the *paxM* promoter were generated by Gibson assembly (2.6.12). These modified genes were based on *E. festucae* Fl1 *perA*, with alterations made to the A1-domain substrate-binding region. These included altering the identity of three predicted substrate-binding residues in the PerA A1-domain to their PerA-ΔR equivalents [residues 6, 7 and 8(-1) in the NRPS A-domain 10AA code; Figure 4.10B] to make plasmid pDB17 (9.3.10), expressing protein PerA-17 (Figure 4.10A & B). Replacement of the entire *perA* sequence encoding the PerA A1-domain substrate-binding region, defined as positions 1211–1722 of *perA* (with position 1 being the first nucleotide of the *perA* start codon), with that of *E. festucae* E2368 *perA-ΔR* generated plasmid pDB31 (9.3.20), expressing protein PerA-31 (Figure 4.10A & B). This substrate-binding region was defined so as to include all codons encoding 10AA code residues, with the exception of the final invariant lysine. Finally, replacement of the entire *perA* sequence encoding the A1-domain substrate-binding region with that of *E. typhina* ssp. *poae* E1022 *perA-ΔR* generated plasmid pDB32 (9.3.21), expressing protein PerA-32 (Figure 4.10A & B). The sequence fidelity of each insert was confirmed by sequencing (2.6.13).

Plasmids pDB17, pDB31, pDB32 and the empty vector plasmid pRS426 were each co-transformed (2.9.2) into wt *P. paxilli* protoplasts with pSF15.15 (9.3.2), which provided hygromycin resistance. Transformations with circular pRS426, which provided the vector backbone for all other constructs, were performed in order to provide experimental negative controls. Crude gDNA was prepared from spore suspensions of each transformant (2.5.6), and pDB17, pDB31 and pDB32 transformants were screened by PCR (2.6.3) using primer pairs KG11/KG12, PPperA_SCR_1/ PPperA_SCR_2 and PPperA_SCR_3/PPperA_SCR_4. These primer pairs amplify the 3’ end of *perA*, from within the 5’ end of *perA* across the upstream *paxM* promoter sequence and from within the 3’ end of *perA* across the downstream *paxM* terminator sequence, respectively. The pRS426 transformants

were screened with primer pair pRS426_SCR_F/pRS426_SCR_R, which amplified within the pRS426 plasmid. Three pRS426 transformants that tested positive for pRS426 by PCR were selected for inclusion in future analyses (Table 4.3). Transformants for which extracted gDNA was successfully PCR amplified with each of the above primer pairs were then grown in liquid culture (2.4.3.1) for 5 days and RNA was extracted (2.7.1) from the isolated mycelia harvested (2.4.4). The RNA samples were used for RT-PCR (2.7.3; 2.6.3) with primer pair KG11/KG12 to confirm *perA* expression and tub2_F/tub2_R to confirm the absence of gDNA contamination (Figure 4.11). Three strains from each transformation that were positive in all PCR/RT-PCR screens were then selected for further experiments (Table 4.3). These transformants were each grown in triplicate [three liquid cultures (2.4.3.1) per transformant, with three transformants per construct] for 5 days at 22°C/200 RPM alongside the *perA* expressing strains 05-31, 05-33 and 05-43. Mycelia were harvested (2.4.4), freeze-dried and peramine was extracted (2.11.1) from and 20 mg of lyophilised mycelia from each sample. The peramine concentration in each extract was determined by LCMS (2.11.3). The results showed that the three amino acid substitutions in PerA-17 resulted in a significant ($P = 0.01$; one-tailed Welch's *t*-test), approx. 13-fold reduction in peramine production (Figure 4.10C). The peramine concentration in PerA-31 expressing transformants was further reduced relative to PerA-17, but this difference was not significant (Figure 4.10C). Finally, peramine was not detected in those transformants expressing PerA-32 (Figure 4.10C).

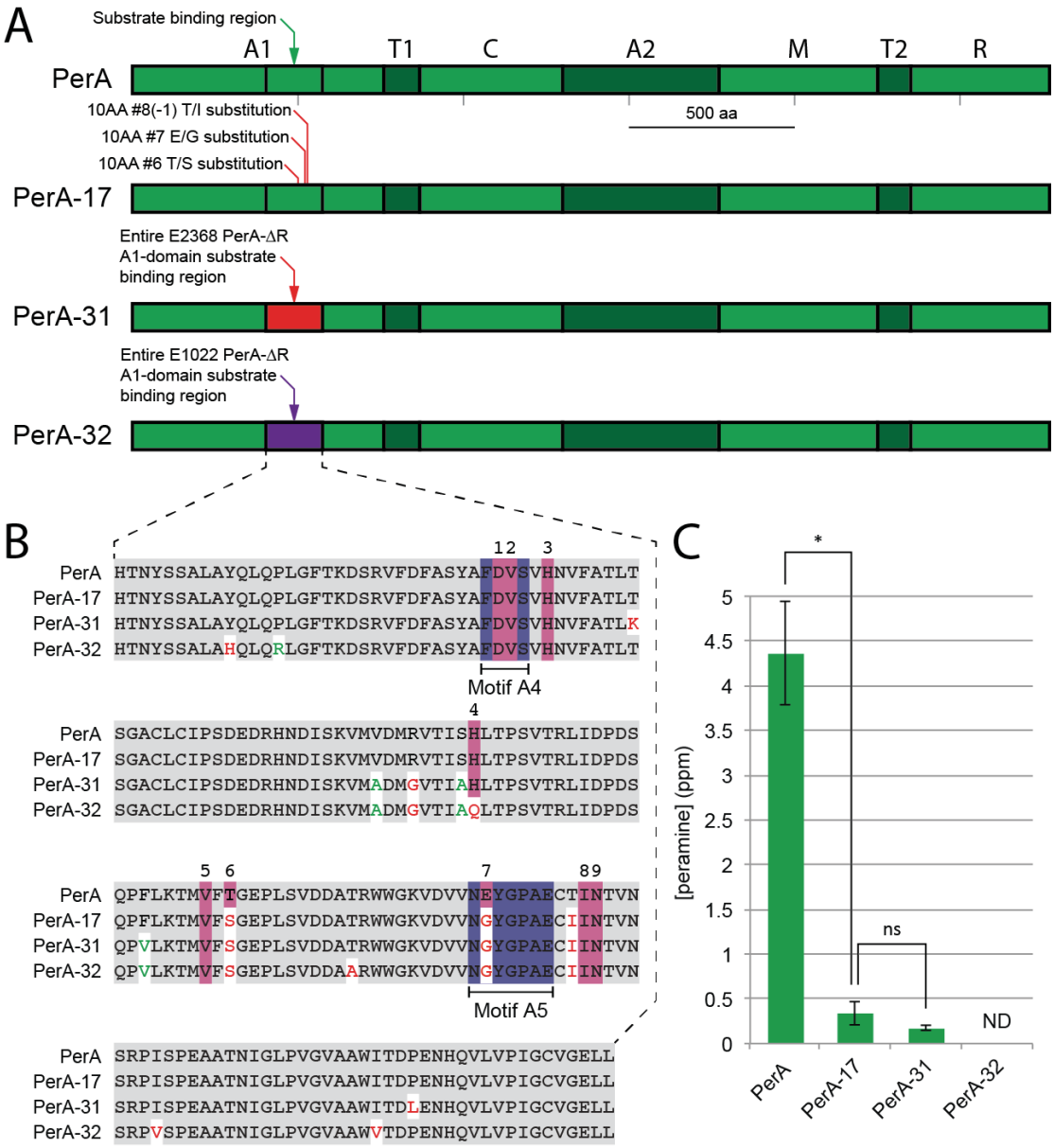


Figure 4.10 Comparison of A1-domain substrate specificity between PerA and PerA-ΔR proteins. (A) Protein sequence map of PerA alongside PerA-17 (3 PerA-ΔR substitutions on or adjacent to 10AA code residues), PerA-31 (A1-domain substrate-binding region swap with *E. festucae* E2368 PerA-ΔR) and PerA-32 (A1-domain substrate-binding region swap with *E. typhina* ssp. *poae* E1022 PerA-ΔR). (B) A1-domain substrate-binding regions from peramine-producing *E. festucae* Fl1 PerA and modified PerA proteins. Red residues are PerA-ΔR unique, while green residues show substitutions found in other peramine producing PerA proteins. (C) Mean peramine concentration of *P. paxilli* strains expressing PerA, PerA-17, PerA-31 or PerA-32. Values show the mean peramine concentration across three independent strains expressing each protein (Table 4.3), with each strain grown in triplicate. Error bars show standard error of the mean. Significance of difference between means was calculated using one-tailed Welch's *t*-test ($P = 0.01$ for PerA vs. PerA-17; $P = 0.17$ for PerA-17 vs. PerA-31). ND = peramine not detected, ns = not significant.

4.4.2 Identifying key changes in PerA-ΔR proteins by domain-swap analysis

Hybrid *perA*/*perA-ΔR* genes under control of the *paxM* promoter were generated by Gibson assembly (2.6.12) to investigate how PerA-ΔR-unique amino acid substitutions alter protein function. These hybrid genes were based on the *E. festucae* E2368 *perA-ΔR* allele. Regions of this allele that showed significant sequence divergence from the *E. festucae* Fl1 *perA* allele were sequentially replaced with homologous fragments of the Fl1 *perA* allele (Figure 4.12). The resulting plasmids were pDB20 (9.3.13), expressing protein PerA-20; pDB29 (9.3.18), expressing protein PerA-29; pDB30 (9.3.19), expressing protein PerA-30; and pDB43 (9.3.25), expressing protein PerA-43 (9.3.25; Figure 4.12). The sequence fidelity of each insert was confirmed by sequencing (2.6.13). The constructs were designed to test whether the replacement of corresponding PerA-ΔR domains with the PerA T2/R-domains (PerA-20), M/T2/R-domains (PerA-29), C/M/T2/R-domains (PerA-30) or C/T2/R-domains (PerA-43) would enable these modified PerA-ΔR proteins to produce peramine. Each of these plasmids was transformed, screened by PCR and screened by RT-PCR as described for e.g. pDB17 transformants in Section 4.4 (Figure 4.11). The selected strains are detailed in Table 4.3. These strains were then grown and analysed for peramine production as described in Section 4.4. Due to time constraints strains 29-1, 30-2, 30-28 and 31-9 have not yet been analysed (Table 4.3). The results showed qualitatively that a minimum of the PerA M, T2 and R-domains, as found in PerA-29, was required to restore peramine production to PerA-ΔR (Figure 4.12). Peramine production from the PerA-29 protein was slightly less than half that of PerA (Figure 4.12), but the lack of a third *perA-29* transformant precluded statistical interrogation of this difference at this stage. PerA-43, which contained the PerA C-domain without the PerA M-domain, did not restore peramine production in two out of three transformants and a peramine signal just above the limit of detection detected for the third (Figure 4.12; Appendix 9.1). Combining the PerA C and M-domains in PerA-30 resulted in levels of peramine production similar to that of PerA, and approx. double that of PerA-29 (Figure 4.12). However, a lack of biological replicates again prevented statistical interrogation of this difference.

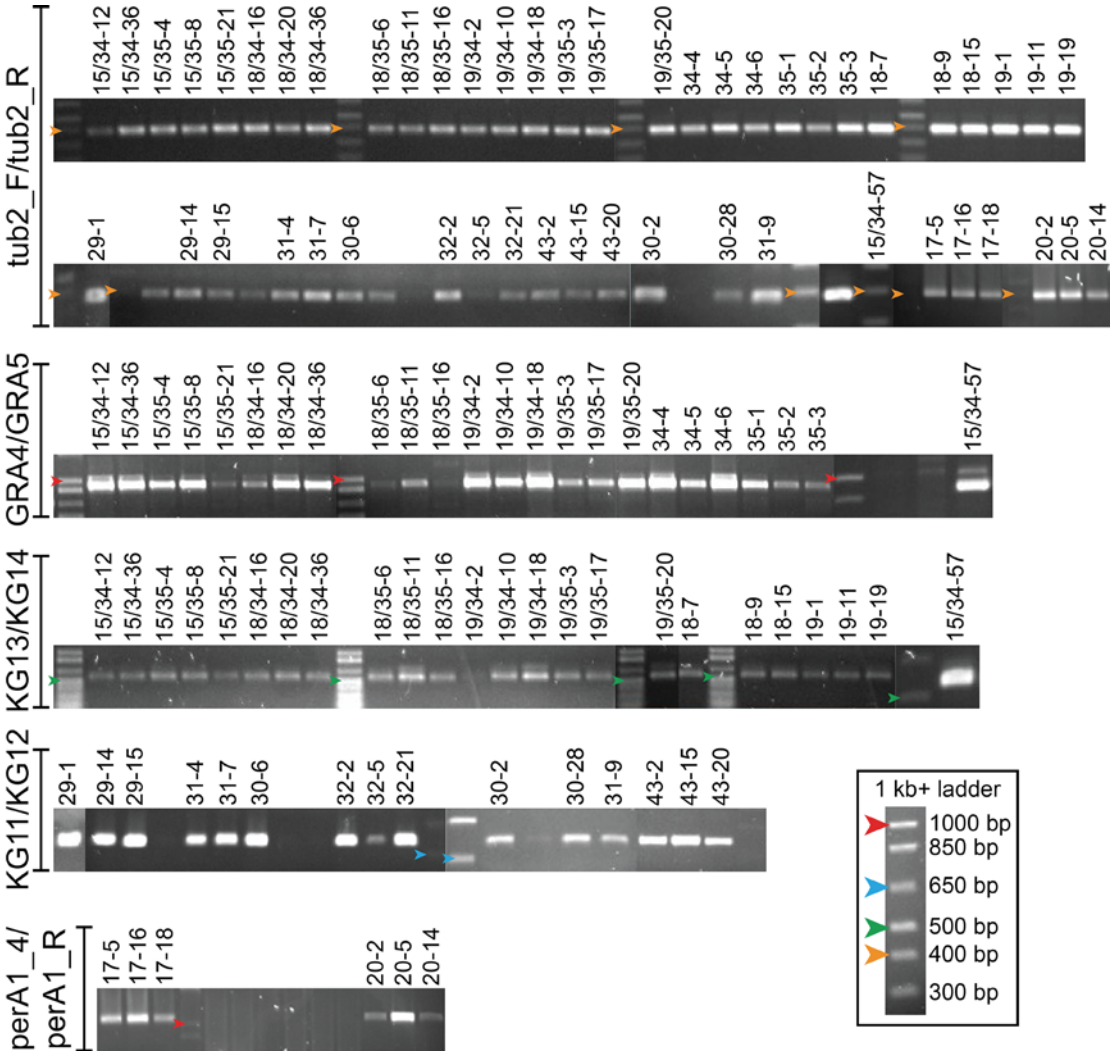


Figure 4.11 RT-PCR screening of *P. paxilli* transformants. Gels showing the results from RT-PCR screening of cDNA from various *P. paxilli* transformants. Primer pair tub2_F/R amplifies across two introns in the *P. paxilli* β-tubulin gene to check for gDNA contamination of the cDNA (expected product size 532 bp for gDNA template and 395 bp for cDNA template). Primer pairs GRA4/GRA5 (product size: 1042 bp gDNA, 940 bp cDNA), KG11/KG12 (product size: 735 bp) and KG13/KG14 (product size: 557 bp) confirmed expression of the *E. festucae* type II thioesterase-encoding gene *tesB* (see Section 4.4.3), *perA* and *perA-ΔR*, respectively. Primer pair PerA1_4/PerA1_R (product size 1073 bp) was used to confirm *perA* expression in pDB17 and pDB20 transformants only. Inset: 1 kb+ ladder (Invitrogen). The 1000 bp, 650 bp and 500 bp bands are indicated on the gel photos using red, blue and green arrowheads, respectively.

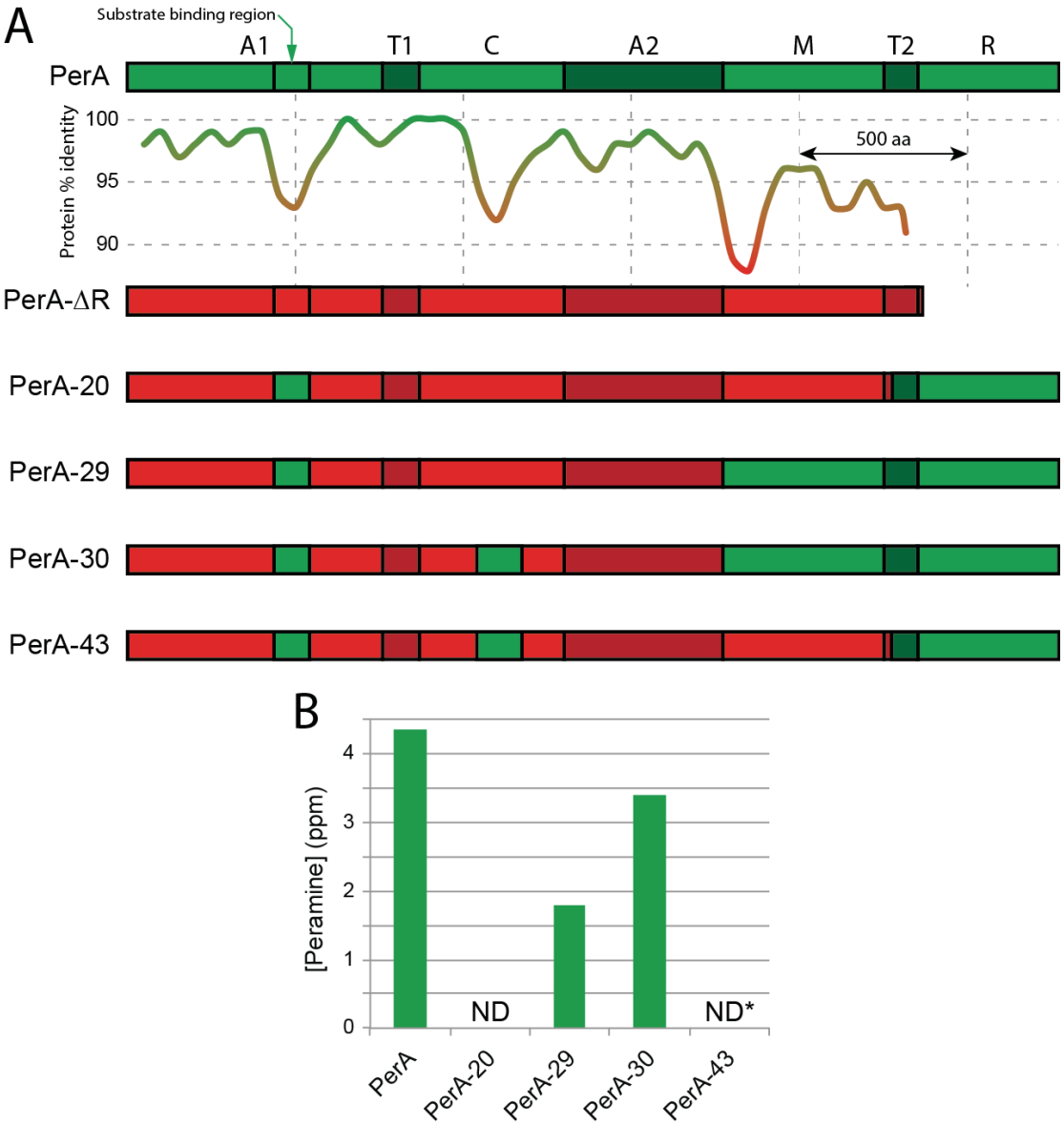


Figure 4.12 Reconstruction of peramine production in PerA-ΔR. (A) Protein sequence maps of *E. festucae* Fl1 PerA, *E. festucae* E2368 PerA-ΔR and the PerA/PerA-ΔR domain-swap hybrid constructs. Regions of the hybrid construct protein sequence maps coloured red originate from E2368 PerA-ΔR, while green coloured regions originate from Fl1 PerA. The graph separating PerA and PerA-ΔR protein sequence maps is a reproduction of the sliding window protein sequence identity analysis described in Figure 4.2. (B) Mean peramine concentration of *P. paxilli* strains expressing PerA (strains 05-31, 05-33, 05-43), PerA-20 (strains 20-2, 20-5, 20-14), PerA-29 (strains 29-14, 29-15), PerA-30 (strain 30-6) or PerA-43 (strains 43-2, 43-15, 43-20). The peramine concentration was averaged across triplicate cultures for each strain. ND = peramine not detected. ND* = peramine not detected in majority of transformants.

Table 4.3 Relevant *P. paxilli* strains for PerA-ΔR A1-domain substrate specificity and domain swap analyses

Gene in *P. paxilli* strain	Strain ID	Plasmid 1	Plasmid 2	Purpose
perA	05-31	pDB05	pAN7-1	peramine positive control
perA	05-33	pDB05	pAN7-1	peramine positive control
perA	05-43	pDB05	pAN7-1	peramine positive control
perA-17	17-5	pDB17	pSF15.15	substrate specificity analysis
perA-17	17-16	pDB17	pSF15.15	substrate specificity analysis
perA-17	17-18	pDB17	pSF15.15	substrate specificity analysis
perA-20	20-2	pDB20	pSF15.15	domain swap analysis
perA-20	20-5	pDB20	pSF15.15	domain swap analysis
perA-20	20-14	pDB20	pSF15.15	domain swap analysis
perA-29	29-1*	pDB29	pSF15.15	domain swap analysis
perA-29	29-14	pDB29	pSF15.15	domain swap analysis
perA-29	29-15	pDB29	pSF15.15	domain swap analysis
perA-30	30-2*	pDB30	pSF15.15	domain swap analysis
perA-30	30-6	pDB30	pSF15.15	domain swap analysis
perA-30	30-28*	pDB30	pSF15.15	domain swap analysis
perA-31	31-4	pDB31	pSF15.15	substrate specificity analysis
perA-31	31-7	pDB31	pSF15.15	substrate specificity analysis
perA-31	31-9*	pDB31	pSF15.15	substrate specificity analysis
perA-32	32-2	pDB32	pSF15.15	substrate specificity analysis
perA-32	32-5	pDB32	pSF15.15	substrate specificity analysis
perA-32	32-21	pDB32	pSF15.15	substrate specificity analysis
perA-43	43-2	pDB43	pSF15.15	domain swap analysis
perA-43	43-15	pDB43	pSF15.15	domain swap analysis
perA-43	43-20	pDB43	pSF15.15	domain swap analysis
-	426-15	pRS426	pSF15.15	peramine negative control
-	426-16	pRS426	pSF15.15	peramine negative control
-	426-17	pRS426	pSF15.15	peramine negative control

* Indicates strains yet to be tested for peramine.

4.4.3 Analysis of product release in PerA-ΔR proteins

In the absence of an R-domain, PerA-ΔR proteins have no obvious mechanism for product release. It was hypothesised that the changes observed in the protein sequence of PerA-ΔR T2-domains might be involved in enabling this release. To investigate this possibility a *E. festucae* Fl1 *perA* gene truncated at the same location as *E. festucae* E2368 *perA-ΔR* was inserted between the *paxM* promoter and

terminator sequences by Gibson assembly (2.6.12). This generated plasmid pDB18 (9.3.11), expressing protein PerA-18 (Figure 4.13). A similar second construct was prepared where the T2-domain-encoding sequence of *perA-18* was replaced with that of E2368 *perA-ΔR* to generate plasmid pDB19 (9.3.12), expressing protein PerA-19 (Figure 4.13). The sequence fidelity of each insert was confirmed by sequencing (2.6.13).

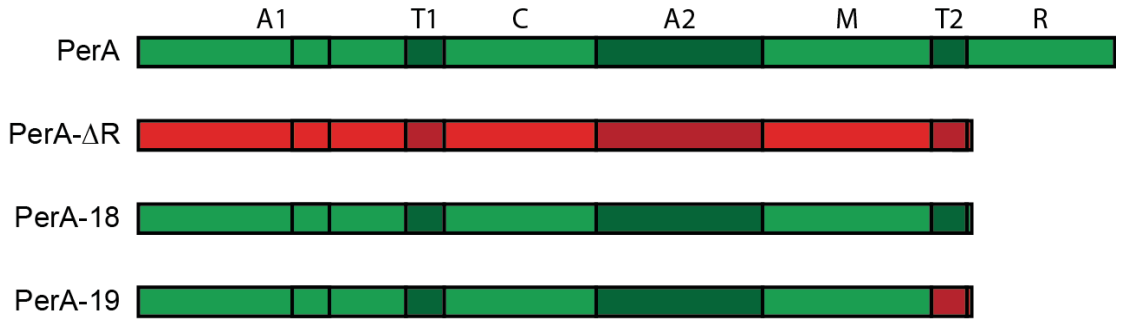


Figure 4.13 Constructs for analysis of Per-ΔR product and product release. The green regions in PerA-18 and PerA-19 originate from *E. festucae* Fl1 PerA, while the red regions originate from *E. festucae* E2368 PerA-ΔR.

The PerA-ΔR-unique substitution at position +24 relative to the predicted T2-domain 4'PP attachment site (Figure 4.4) was identified as indicating a change in the protein or domain partner(s) with which this domain interacts (Zhou et al., 2006; David Ackerley, personal communication). It was also observed that many of the PerA-ΔR changes in the T2-domain occurred within regions previously shown to interact with NRPS-maintaining type II thioesterase (TEII) proteins, specifically the loop containing the 4'PP attachment site, the N-terminus of α-helix II and the region containing α-helix III (Koglin et al., 2008; Koglin et al., 2006). TEIIs have the demonstrated capacity to release amino acids, and even tripeptides, bound via thioester bonds to the phosphopantetheine moiety of holo-T-domains (Koglin et al., 2008; Schwarzer et al., 2002). I therefore hypothesised that the PerA-ΔR-unique T2-domain substitutions might increase the affinity for a TEII protein, providing a product release mechanism for PerA-ΔR proteins. As PerA-ΔR proteins would have co-evolved with the *Epichloë* TEII protein, not the *P. paxilli* TEII protein, it was

decided to co-express *perA-18* and *perA-19* with the *E. festucae* TEII-encoding gene. There were two TEII-encoding genes annotated on the *E. festucae* E2368 genome (Schardl et al., 2013a), however one of these straddled a gap between two contigs on a larger scaffold assembly. The homologous sequences from the *E. festucae* Fl1 genome (gene models EfM3.012870 and EfM3.113290) were therefore analysed further.

TEIIs in fungi have yet to be characterised and extremely low conservation between characterised bacterial TEIIs and fungal TEIIs made determining which of these *E. festucae* TEIIs fulfils the NRPS/PKS maintenance role difficult. Primary structure conservation between the two predicted *E. festucae* TEII proteins was low (21%), indicating they were not the result of a recent gene duplication event. BLASTn searches found that homologues of both genes were present across *Epichloë.* However, BLASTn analysis across the Clavicipitaceae found that homologs to EfM3.113290 were present in the genomes of all 11 species surveyed outside of *Epichloë*, whereas EfM3.012870 homologues were not identified in the genomes of *Atkinsonella hypoxylon*, *Atkinsonella texensis*, *Balansia obtecta* and *Claviceps paspali* (Table 4.4). BLASTp searches against genomes of notable model Ascomycota species known to contain NRPS genes were also performed. The results revealed that conservation between the predicted protein encoded by EfM3.113290 and homologous protein sequences from these model species was consistently higher than between the EfM3.012870 predicted protein product and its closest homologs (Table 4.4).

Based on these observations it was hypothesised that the protein encoded by EfM3.113290, which was named TesB, was the *E. festucae* TEII responsible for maintaining NRPS and PKS enzymes. This was based on the hypothesis that such a TEII would be well conserved across the filamentous fungi, as its role is essential to a huge number of secondary metabolite pathways. The predicted start and stop codons annotated for *tesB* were interrogated by aligning this locus across multiple *Epichloë* genomes, and the results strongly supported the original annotations (Appendix 9.1). Due to the potential for co-evolution between PerA-ΔR and TesB it was desired to co-express *perA-ΔR* with a *tesB* gene sourced from the same isolate (*E. festucae* E2368). However, E2368 is a multi-parental strain that was generated by

backcrossing the F1 progeny of an *E. festucae* E189 × *E. festucae* E434 cross against E189 (Schardl et al., 2013a). E189 is the parental strain of *E. festucae* E2368 from which the *perA-ΔR* gene is inherited (Berry et al., 2015b). The genome sequence of E189 was not available, and it was not known if the E2368 *tesB* originated from E189

Table 4.4 BLAST search results using *E. festucae* Fl1 TEII queries

	Strongest E-value vs. EfM3.11290	**Strongest E-value vs. EfM3.012870**
Clavicipitaceae	**Using BLASTn with CDS query**	**Using BLASTn with CDS query**
Aciculosporium take	3×10^{-37}	2×10^{-26}
Atkinsonella hypoxylon	1×10^{-66}	0.53
Atkinsonella texensis	4×10^{-72}	0.42
Balansia obtecta	6×10^{-65}	0.46
Beauveria bassiana	1×10^{-144}	2×10^{-108}
Claviceps fusiformis	8×10^{-41}	1×10^{-101}
Claviceps paspali	3×10^{-54}	0.11
Claviceps purpurea	2×10^{-24}	1×10^{-81}
Metarhizium acridum	2×10^{-50}	1×10^{-23}
Metarhizium robertsii	2×10^{-50}	8×10^{-22}
Periglandula ipomoeae	2×10^{-59}	1×10^{-143}
Ascomycota	**Using BLASTp with predicted protein query**	**Using BLASTp with predicted protein query**
Aspergillus flavus	7×10^{-108}	2×10^{-34}
Aspergillus nidulans	4×10^{-112}	3×10^{-33}
Fusarium graminearum	1×10^{-148}	3×10^{-138}
Neurospora crassa	4×10^{-139}	no matches

or E434. Primers were therefore designed against E2368 *tesB* and were successfully used to PCR amplify (2.6.4) the E189 homolog. This PCR fragment was then cloned between the *tefA* promoter sequence from *Aureobasidium pullulans* (Vanden Wymelenberg et al., 1997) and terminator sequence from the *P. paxilli tesB* homolog (Berry et al., 2015a) to generate pDB34 (9.3.22), and between the promoter and terminator sequences of the *P. paxilli tesB* homolog to generate pDB35 (9.3.23). The *A. pullulans tefA* promoter sequence has been shown to initiate strong, constitutive expression of genes placed downstream in *E. festucae* despite its heterologous

origins (Takemoto et al., 2006), and was considered likely to do the same in *P. paxilli*. Previous studies had shown that increasing the concentration of a TEII *in vitro* caused a corresponding increase in tripeptide product release for an NRPS with no product release domain (Schwarzer et al., 2002). A similar effect was hypothesised for overexpression of *tesB* from E189 *in vivo*, potentially increasing product release. An *hph* expression cassette that conferred hygromycin resistance was included in both pDB34 and pDB35 so that co-transformations could be done with *perA-ΔR*-expressing plasmids without requiring a third plasmid to confer antibiotic resistance for selection. The sequence fidelity of these plasmids was confirmed by comprehensive sequencing (2.6.13), and this confirmed that the E2368 and E189 *tesB* genes are identical.

P. paxilli protoplasts were transformed with pDB18, pDB19, pDB18/pDB34, pDB18/pDB35, pDB19/pDB34, pDB19/pDB35, pDB34 or pDB35. The two latter transformations were designed to provide background control strains for metabolite analyses of strains expressing P*tefA-tesB* and P*tesB-tesB*, respectively. Crude gDNA was prepared from spore suspensions of each transformant (2.5.6), with gDNA from pDB18 and pDB19 transformants screened for PCR amplification (2.6.3) using primer pairs p9f/p9r, PPperA_SCR_1/PPperA_SCR_2 and perA3_3/PPperA_SCR_4. These primer pairs were designed for amplification of the A2-domain-encoding region of *perA-ΔR*, from within *perA-ΔR* across P*paxM* and from within *perA-ΔR* across T*paxM*, respectively (9.3.11; 9.3.12). Crude gDNA from transformants containing pDB34 or pDB35 was screened for PCR amplification using primer pairs GRA4/GRA5 and KG5/KG6, which were designed for amplification within the *E. festucae tesB* gene and within the terminator sequence of the *P. paxilli tesB* homolog, respectively (9.3.22; 9.3.23). Crude gDNA from transformants containing pDB34 was additionally screened for PCR amplification using primer pair KG8/KG2, while gDNA from transformants containing pDB35 was screened for PCR amplification using primer pair KG1/KG2. These primer pairs were designed for PCR amplification from within the *tesB* gene across the P*tefA* and *P. paxilli* P*tesB* sequences, respectively. Transformants from which gDNA was successfully used as template for PCR amplification using all primer pairs were then grown in liquid culture (2.4.3.1), harvested (2.4.4), then RNA was extracted from the isolated mycelia (2.7.1). These

RNA samples were used as template for RT-PCR (2.7.3; 2.6.3) using primer pair KG13/KG14, which were designed for PCR amplification of the 3' end of *perA-ΔR* to confirm *perA-ΔR* expression (Figure 4.11). Expression of the *tesB* was similarly confirmed by RT-PCR using primer pair GRA4/GRA5, which were designed for PCR amplification of most of the E189 *tesB* gene (Figure 4.11). Three strains from each transformation that were positive in all PCR/RT-PCR screens were then selected for further experiments (Table 4.5). Each *P. paxilli* 18, 19, 18/34, 18/35, 19/34, 19/35, 426, 34 and 35 strain was grown in triplicate (2.4.3.1). Mycelia were then harvested from each culture (2.4.4), freeze-dried and metabolite extractions were performed on 100 mg of lyophilised mycelia from each sample (2.11.1). These extracts were lyophilised and resuspended in 50 µL of 50% methanol to concentrate the samples (v/v). Unfortunately, due to equipment failures, analysis of these samples has yet to be performed.

4.4.4 Identification of novel PerA-ΔR products

The *perA-ΔR* gene from *E. festucae* E2368 was placed between the *paxM* promoter and terminator sequences by Gibson assembly to generate plasmid pDB15 (9.3.9). *P. paxilli* protoplasts were then transformed with pDB15, pDB15/pDB34 or pDB15/pDB35. The *tesB*-expressing pDB34 and pDB35 constructs were included as TesB was hypothesised to catalyse product release from PerA-ΔR, and might therefore enable or increase product generation. Transformants were then screened as described in Section 4.4.3, with the same primers used for screening *perA-ΔR* as were used for *perA-18* and *perA-19*. Selected strains are listed in Table 4.5. Each *P. paxilli* 15, 15/34, 15/35, 426, 34 and 35 strain was grown in triplicate (2.4.3.1). Mycelia were then harvested from each culture (2.4.4), freeze-dried and metabolite extractions were performed on 100 mg of lyophilised mycelia from each sample (2.11.1). These extracts were lyophilised and resuspended in 50 µL of 50% methanol to concentrate the samples (v/v). Unfortunately, due to equipment failures, analysis of these samples has yet to be performed.

Table 4.5 Relevant *P. paxilli* strains for PerA-ΔR product release and novel product identification analyses

Gene(s) in *P. paxilli* strain	Strain ID	Plasm-id 1	Plasmid 2	Purpose
perA	05-31	pDB05	pAN7-1	peramine positive control
perA	05-33	pDB05	pAN7-1	peramine positive control
perA	05-43	pDB05	pAN7-1	peramine positive control
perA-ΔR	15-2	pDB15	pSF15.15	novel product identification
perA-ΔR	15-4	pDB15	pSF15.15	novel product identification
perA-ΔR	15-5	pDB15	pSF15.15	novel product identification
perA-ΔR/P*tefA-tesB*	15/34-12	pDB15	pDB34	novel product identification
perA-ΔR/P*tefA-tesB*	15/34-36	pDB15	pDB34	novel product identification
perA-ΔR/P*tefA-tesB*	15/34-57	pDB15	pDB34	novel product identification
perA-ΔR/P*tesB-tesB*	15/35-4	pDB15	pDB35	novel product identification
perA-ΔR/P*tesB-tesB*	15/35-8	pDB15	pDB35	novel product identification
perA-ΔR/P*tesB-tesB*	15/35-21	pDB15	pDB35	novel product identification
perA-18	18-7	pDB18	pSF15.15	product release analysis
perA-18	18-9	pDB18	pSF15.15	product release analysis
perA-18	18-15	pDB18	pSF15.15	product release analysis
perA-18/P*tefA-tesB*	18/34-16	pDB18	pDB34	product release analysis
perA-18/P*tefA-tesB*	18/34-20	pDB18	pDB34	product release analysis
perA-18/P*tefA-tesB*	18/34-36	pDB18	pDB34	product release analysis
perA-18/P*tesB-tesB*	18/35-6	pDB18	pDB35	product release analysis
perA-18/P*tesB-tesB*	18/35-11	pDB18	pDB35	product release analysis
perA-18/P*tesB-tesB*	18/35-16	pDB18	pDB35	product release analysis
perA-19	19-1	pDB19	pSF15.15	product release analysis
perA-19	19-11	pDB19	pSF15.15	product release analysis
perA-19	19-19	pDB19	pSF15.15	product release analysis
perA-19/P*tefA-tesB*	19/34-2	pDB19	pDB34	product release analysis
perA-19/P*tefA-tesB*	19/34-10	pDB19	pDB34	product release analysis
perA-19/P*tefA-tesB*	19/34-18	pDB19	pDB34	product release analysis
perA-19/P*tesB-tesB*	19/35-3	pDB19	pDB35	product release analysis
perA-19/P*tesB-tesB*	19/35-17	pDB19	pDB35	product release analysis
perA-19/P*tesB-tesB*	19/35-20	pDB19	pDB35	product release analysis
P*tefA-tesB*	34-4	pDB34	-	negative control
P*tefA-tesB*	34-5	pDB34	-	negative control
P*tefA-tesB*	34-6	pDB34	-	negative control
P*tesB-tesB*	35-1	pDB35	-	negative control
P*tesB-tesB*	35-2	pDB35	-	negative control
P*tesB-tesB*	35-3	pDB35	-	negative control
-	426-15	pRS426	pSF15.15	negative control
-	426-16	pRS426	pSF15.15	negative control
-	426-17	pRS426	pSF15.15	negative control

4.5 Discussion

The results presented in this chapter describe genetic and experimental evidence for the function of PerA-ΔR proteins, and provide insight into the evolution of NRPS proteins. Establishment of a heterologous expression system in *P. paxilli* allowed rapid analysis of a large number of different PerA/PerA-ΔR hybrid proteins. Heterologous expression avoided the complications of using *Epichloë festucae*, which requires *in planta* assays to test for metabolite production due to plant-specific expression of SM genes. In addition, the plant metabolic background, slow culture growth and minimal sporulation in culture were also avoided. Heterologous expression of *perA* alone is sufficient to enable *Penicillium paxilli* to produce peramine, which supports the hypothesis of Tanaka et al. (2005) that PerA is the only pathway-specific protein required for peramine production. It should be noted, however, that the PerA proteins require post-translational modification by 4'phosphopanteteinyl transferases (4'-PPTase) before they become functional (Lambalot et al., 1996; Reuter et al., 1999; Walsh et al., 1997), and would also require ongoing maintenance from a type II thioesterase (TEII), which are required to deacetylate acyl-4'PP moieties originating from acetyl-CoA donors and remove aberrant amino acid substrates from mis-primed T-domains (Schwarzer et al., 2002). The fact that peramine production in *P. paxilli* was successful therefore shows that the native *P. paxilli* 4'-PPTase—and likely also the native TEII protein—are functional, and reinforces their generic nature.

The *paxM* gene, which provided the promoter for *perA* expression in *P. paxilli*, encodes a protein required for the first committed step in paxilline biosynthesis (Saikia et al., 2006). Therefore it was not surprising to see that peramine production in *P. paxilli* was only detected after 72 h of growth, as this is also the first time point at which paxilline is detected (Telfer, 2000). However, unlike paxilline, the maximum peramine concentration was achieved after only 96 h. Paxilline continues to accumulate until at least 144 h, and possibly even longer (Telfer, 2000). As the *paxM* promoter is intimately involved with paxilline production, it is unlikely that transcriptional changes cause this abrupt cut-off in peramine production. It has been previously shown that the dry mass of *P. paxilli*

peaks then decreases after approx. 48 h of growth, with the energy accumulated in biomass thought to be redirected towards paxilline production (Telfer, 2000). Expression of *perA* in a Δ*PAX*-cluster strain showed that eliminating the pax cluster genes, and presumably the associated metabolic cost of paxilline production, did not increase peramine production. Nevertheless, given that *P. paxilli* growth in culture peaks by 48 h (Telfer, 2000), limitation of peramine production due to substrate starvation is a likely explanation for why peramine production stops after 96 h. Addition of arginine, a predicted PerA substrate (Tanaka et al., 2005), and proline, which is readily converted to the predicted PerA substrate 1-pyrroline-5-carboxylate (P5C; Aral and Kamoun, 1997), to the growth media before inoculation did not improve peramine production. This might be because the increased proline concentration did not translate to increased P5C concentration, though it seems more likely that these amino acids in excess were converted to other compounds prior to *perA* induction. Placing *perA* under the control of a strong promoter that induces transcription earlier during growth might enable increased peramine production due to increased substrate availability. However, this might also have unintended negative consequences on *P. paxilli* primary metabolism and growth, and it was for this reason that a SM-specific promoter was originally selected. An alternative might be to feed the cultures appropriate substrates after *perA* induction to increase peramine production.

Comparison of the A1-domain substrate-binding sequences between PerA and PerA-ΔR proteins identified a number of PerA-ΔR-unique substitutions to the predicted substrate-binding residues. The number of substitutions observed is consistent with other NRPS A-domains for which substrate specificity changes have been documented. One example is the three-module NRPS LpsA from the ergot alkaloid biosynthesis pathway. The substrate specificity of each LpsA A-domain varies considerably across the Clavicipitaceae (Young et al., 2015). This diversity is documented in Table 4.6, showing how only a few binding residue substitutions are required to alter substrate specificity. The A1-domain predicted substrate-binding residues also varied somewhat between PerA-ΔR proteins, with some lineages containing additional substitutions at these positions. This suggests divergent selection on these positions has been maintained after radiation from the last

common ancestor, suggesting selective pressure to further alter or refine the PerA-ΔR A-domain substrate specificity. This hypothesis is supported by the observation that when the A1-domain substrate-binding region of PerA was replaced with that of *E. festucae* E2368 PerA-ΔR, the resulting PerA-31 protein retained the capacity to produce small but detectable levels of peramine. In contrast, use of the A1-domain substrate-binding region from *E. typhina* ssp. *poae* E1022 PerA-ΔR, which contains additional substrate-binding residue substitutions relative to E2368 PerA-ΔR, completely abolished peramine production for PerA-32.

Table 4.6 Substrate diversity in LpsA proteins[1]

Host species	Protein	Module #	8AA code[2]	Substrate[1]
E. festucae	LpsA	1	IFFCGGPL	Ala
C. purpurea	LpsA1	1	AIFCGGPL	Ala
C. purpurea	LpsA2	1	AVFCVGPA	Val
P. ipomoeae	LpsA	1	LFFCGGPL	Ala
E. festucae	LpsA	2	AVMVAAVI	Val
C. purpurea	LpsA1	2	LVGMAAVG	Phe
C. purpurea	LpsA2	2	LAGMGAMI	Leu
P. ipomoeae	LpsA	2	LVAMGAVI	Leu
E. festucae	LpsA	3	ITLVAGLV	Pro
C. purpurea	LpsA1	3	ITLVAGLI	Pro
C. purpurea	LpsA2	3	ITLVAGLI	Pro
P. ipomoeae	LpsA	3	IMIVAGVI	Ala

[1] Substrates are listed as reported by Young et al. (2015).
[2] 10AA codes were determined by NRPSpredictor2 (Roettig et al., 2011), with the 8AA codes shown here (same as 10AA code but with invariant D and K residues removed).

It is important to note that strains transformed with expression constructs for PerA-ΔR or modified PerA proteins that cannot produce peramine lack an intrinsic positive control (peramine production) to confirm the strain in question has produced significant quantities of full length PerA/PerA-ΔR protein. However, for strains expressing native or modified PerA proteins capable of producing peramine, there was a strong correlation between RT-PCR results indicating *perA* expression and detection of peramine in culture. This result is perhaps surprising, as integration of a plasmid into a host genome requires a change in topology from circular to linear DNA, requiring a breakpoint within the circular plasmid. Given that the *perA* alleles

represent more than half of the total DNA sequence of these plasmids, this breakpoint should frequently occur within the *perA* alleles themselves, disrupting the coding sequence. The fact that this does not seem to occur indicates that either the breakpoint site is non-random, or more likely that the transformed plasmids are integrating in tandem, resulting in multiple intact *perA* expression cassettes. Such tandem integrations for circular plasmid are common in *E. festucae* (Y. Lukito, personal communication), indicating the same is likely true for *P. paxilli*. Positive results from RT-PCR screens for expression of transformed *perA* alleles are therefore likely to be accurate for predicting functional protein production. Determining *perA* allele copy number in several *perA*-expressing strains by qPCR could confirm this tandem insertion hypothesis.

A-domain substrate predictions from protein primary structures are now reasonably accurate for bacterial NRPS sequences (Roettig et al., 2011). However, *in silico* predictions for fungal NRPSs remain relatively unreliable (Kalb et al., 2013). This is due to the lower number of fungal A-domain sequences with characterised substrates compared to bacteria, and subtle differences to binding site architecture that mean bacterial substrate-specificity codes cannot be directly applied to fungal A-domains (Kalb et al., 2013; Lee et al., 2010). This is evidenced when predicting the A1-domain substrates of PerA and PerA-ΔR proteins using NRPSpredictor2 (Roettig et al., 2011). NRPSpredictor2 does indicate a change in A1-domain substrate from α-aminoadipic acid for PerA to Val/Leu for PerA-ΔR proteins. However, the fact that the PerA A1-domain substrate is incorrectly identified as α-aminoadipic acid raises serious doubts about the validity the PerA-ΔR assignments. It is important that any future attempts to express and purify the PerA A1-domain for study *in vitro* for substrate or structural determination take into account the 250 aa N-terminal extension with predicted structural conservation to the second subunit of an NRPS C-domain. The recently published study by Kalb et al. (2015) showed that this domain extension is required for substrate AMPylation by the downstream A-domain, and the same is likely true for PerA and PerA-ΔR proteins. The presence of this domain extension suggests that the ancestral sequence of the PerA A1-domain may have been part of an elongation module instead of its current position within an initiation module. The ancestral A-domain would therefore have physically

interacted with an upstream C-domain, and this interaction might have been required to maintain the A-domain structural integrity even after relocation to an initiation module, as proposed by (Kalb et al., 2015).

Comparison of the PerA/PerA-ΔR domain swap analysis proteins PerA-20 and PerA-29 showed that the PerA M-domain is required to restore peramine production to PerA-ΔR. As there are no changes to PerA-ΔR A2-domains that indicate a change in substrate, this suggests that the PerA-ΔR second module is still likely to incorporate arginine into any dipeptide product, but with an altered methylation state relative to PerA. This might consist of total loss of arginine methylation, or methylation of an atom other than the arginine α-amino nitrogen. The specific localisation of PerA-ΔR-unique residue substitutions to the N-terminus of the M-domain suggests that this M-domain functional change is specific to all PerA-ΔR proteins. However, this cannot be conclusively determined from the experiments presented here as only PerA-ΔR from *E. festucae* E2368 was tested. Testing of an *E. bromicola*/*E. typhina* PerA-ΔR M-domain, or artificial introduction of only the PerA-ΔR-unique residues into an otherwise functional PerA protein, could resolve this.

Analysis of protein PerA-43 showed that reintroduction of the PerA C-domain, in addition to the PerA T2 and R-domains, was not sufficient to restore peramine production. Comparison of peramine production by PerA-29 and PerA-30 suggests that the PerA C-domain may increase peramine production relative to the PerA-ΔR C-domain by a factor of approx. two; however, analysis of additional samples is required to determine the significance of this result. Nevertheless, while NRPS substrate specificity is primarily mediated by A-domains (Eisfeld, 2009; Hur et al., 2012), previous studies have shown that NRPS C-domains exhibit strict stereoselectivity and can also exhibit side-chain selectivity towards the incoming substrate (Ehmann et al., 2000; Lautru and Challis, 2004). The hyper-variability and apparently lowered peramine production efficiency observed in the *E. festucae* E2368 PerA-ΔR C-domain relative to *E. festucae* Fl1 PerA might therefore be explained by evolution towards increased binding affinity for the PerA-ΔR A1-domain substrate.

Unlike neofunctionalisation events driven by gene duplications, such as is observed for the NRPS-encoding *lpsA1* and *lpsA2* genes in *Claviceps purpurea* (Young

et al., 2015), *perA-ΔR* appears to simply be an alternative allele of *perA*. This is evidenced by the fact that *perA-ΔR* appears to be at the same genetic locus as *perA*, although the large transposon relics often associated with *perA-ΔR* alleles can obscure locus determination (Berry, 2011; Fleetwood et al., 2011; Schardl et al., 2013a). All other peramine-negative *perA* alleles identified to date, which usually contain small frameshift-causing indels or larger deletions within *perA*, are only present within a single strain or closely related group of isolates (Chapter 3; Berry et al., 2015b), highlighting the unusually wide distribution of *perA-ΔR* allele. The initial event in the evolution of *perA-ΔR* seems likely to be deletion of the sequence encoding the R-domain through integration of the miniature inverted-repeat transposable element EFT-3m into the end of a *perA* gene, which would have abolished peramine production (Berry et al., 2015b; Fleetwood et al., 2011; Schardl et al., 2013a). It is possible that even in this early form the *perA-ΔR* allele, which likely closely resembled the *perA-18* gene created during this study, may have still provided a selective advantage to its host. This hypothesis is supported by the continued maintenance of *perA-ΔR* in *E. baconii* despite these alleles not encoding the changes to A1-domain substrate-binding residues observed in other PerA-ΔR proteins (Berry et al., 2015b). This selective advantage might have been provided by a P5C-arginine dipeptide similar to peramine, but without the intra-molecular cyclisation observed in peramine. Alternatively, a minor product of the ancestral PerA-ΔR protein might have provided this advantage; it has been previously observed that NRPS A-domains are sometimes also able to bind secondary substrates in addition to their primary amino acid substrate (Rausch et al., 2005). It is possible that the PerA A1-domain has such a secondary substrate, and that deletion of the R-domain resulted in production of an advantageous dipeptide that incorporated this secondary substrate. This could have provided the selective pressure that maintained the *perA-ΔR* allele following the R-domain deletion. This would also have applied diversifying selection on the A1-domain substrate-binding residues towards increasing the binding efficiency for this secondary substrate, resulting in the substitutions observed today in PerA-ΔR proteins.

The absence of a C-terminal R-domain means that PerA-ΔR proteins lack an integrated domain for product release. It is proposed that the *Epichloë* TEII normally

responsible for NRPS maintenance may fulfil this product release role. Product release from NRPS/PKS megasynthases by TEIIs *in trans* has previously been observed in several systems (Du and Lou, 2010; Kotowska and Pawlik, 2014). Release of a tripeptide product from an NRPS lacking a termination domain by a TEII has also been previously demonstrated *in vitro* (Schwarzer et al., 2002). PerA-ΔR proteins have also accumulated mutations within regions previously shown to interact with TEIIs (Koglin et al., 2008; Koglin et al., 2006). The modified proteins PerA-18 and PerA-19 were designed to investigate product release by PerA-ΔR proteins. These proteins should bind the known PerA substrates (Tanaka et al., 2005), making product identification much easier. Comparing the efficiency with which these proteins release product, particularly in the presence of the *E. festucae* E2368 TEII, may provide significant insight into any release mechanism employed by PerA-ΔR proteins. An alternative hypothesis is that PerA-ΔR proteins interact with a downstream NRPS to produce a longer peptide product. This was considered less likely given that it requires co-evolution of PerA-ΔR with a second megasynthase. PerA-ΔR also does not appear to contain a COM-domain, which would usually mediate interactions between subunits of a larger NRPS complex (Hahn and Stachelhaus, 2004). Indeed several *perA-ΔR* alleles from *E. bromicola* and *E. typhina* have SNPs that introduce stop codons immediately downstream of the T2-domain-encoding sequence, eliminating the translation of what may be superfluous residues for a PerA-ΔR protein.

Phylogenetic analysis of the T2-domain-encoding regions from a variety of *perA* and *perA-ΔR* alleles found that *perA-ΔR* alleles exhibit apparent trans-species polymorphism (TSP; Chapter 3; Berry et al., 2015b). TSP occurs when balancing selection maintains two or more alleles across speciation events such that a specific allele is more closely related to the same allele from a related species than to a different allele from the same species. The suggestion that *perA*/*perA-ΔR* alleles exhibit TSP therefore implies balancing selection to maintain both alleles within a population. This is intriguing not only because it implies that PerA-ΔR proteins confer a selective advantage, but because previously characterised examples of TSP have been dominated by immune system genes (Tesicky and Vinkler, 2015). It is interesting to think of the bio-protective SM of *Epichloë* spp. as analogous to an

innate immune system. Endophyte-grass associations come under pressure from a variety of different vertebrate and invertebrate herbivores, and production of specific secondary metabolites confers resistance or immunity to consumption by some, but not all, of these herbivores. A classic example of this is peramine, which is a strong deterrent against herbivory by adult Argentine stem weevils (*Listronotus bonariensis*; Rowan et al., 1990), but provides no resistance against adult black beetles (Ball et al., 1997). Immune system alleles that exhibit TSP tend to display a similar pattern, where specific alleles provide resistance to certain pathogens, meaning that a particular allele may become dominant due to being advantageous in a particular pathogenic background (Sommer, 2005; Tesicky and Vinkler, 2015). However, changes to that background driven by the introduction of new pathogens and/or the tendency for pathogen evolution to be driven by the dominant host genotype (Sommer, 2005) provides balancing selection that prevents allele fixation. This immune system analogy would suggest that the proteins encoded by *perA* and *perA-ΔR* produce secondary metabolites that provide resistance to different herbivores. Interestingly, the commercial endophyte strain AR37 contains a *perA-ΔR* allele that is identical in sequence to the *E. festucae* E2368 *perA-ΔR* allele analysed in this study. AR37 is an enigmatic strain, as it provides strong resistance to several agriculturally relevant insect pests, but does not produce any of the previously characterised anti-insect peramine or loline secondary metabolites, and does not provide resistance to adult Argentine stem weevils (Johnson et al., 2013a). It has been proposed that production of the indole-diterpenoid epoxy-janthitrems by this strain contributes to this insect resistance (Johnson et al., 2013a). However, these epoxy-janthitrems do not appear to protect against all of the insect species that AR37 confers resistance to (Alison Popay, personal communication).

The TSP hypothesis for the relationship between *perA* and *perA-ΔR* alleles is complicated by the observation that the phylogenetic topology of A2-domain-encoding sequences from *perA-ΔR* alleles in *E. baconii*, *E. bromicola* and *E. festucae* roughly follows species phylogeny (Berry et al., 2015b). This contrasts the TSP-like phylogeny of the T2-domain (Berry et al., 2015b), and is shown to be the result of recombination between *perA* and *perA-ΔR* alelles. Based on the A2-domain phylogeny (Berry et al., 2015b), it appears that independent recombination events

between *perA* and *perA-ΔR* alleles have occurred at similar locations in an ancestor of the *E. bromicola* clade, as well as an ancestor of the *E. festucae*/*E. baconii* clade. These events resulted in hybrid alleles, with the *perA-ΔR* M, T2 and R-domain deletions being introduced into a *perA* allele in the *E. festucae*/*E. baconii* clade, while the *E. bromicola perA* allele additionally obtained the *perA-ΔR* A1-domain substrate-binding region. A second recombination event in the *E. festucae* lineage would then explain why *E. festucae* PerA-ΔR A1-domain substrate-binding regions contain the PerA-ΔR-unique changes, while the *E. baconii* PerA-ΔR proteins do not. There are several *perA-ΔR* alleles within the *E. typhina* complex that contain large deletions (Berry et al., 2015b). If the ancestral *perA-ΔR* sequence of *E. baconii*, *E. bromicola* and *E. festucae* contained a similar genetically disruptive mutation, then recombination with a functional *perA* allele may have provided a repair mechanism to restore PerA-ΔR function. It is not clear if the ancestral *perA-ΔR* allele from before these recombination events remains within the populations of *E. baconii*, *E. bromicola* or *E. festucae*, or if this allele has now been lost. Should such a sequence ever be identified, it would provide significant insight into the mechanisms that have resulted in the *perA-ΔR* alleles described here.

The *E. baconii*/*E. bromicola*/*E. festucae perA-ΔR* alleles are also not the only known example of recombination between *perA* alleles; one of the *perA* alleles of the hybrid *Epichloë* sp. FaTG-2 G3 shows evidence of recombination between a *E. festucae*-type *perA* and *perA-ΔR* alleles (Berry et al., 2015b). This resulted in the introduction of the PerA-ΔR M-domain changes, but not the R-domain deletion, into the protein expressed by this isolate. This would have abolished peramine production by the encoded PerA protein, which might explain why the gene encoding this protein has acquired an early premature stop codon (Berry et al., 2015b). Taken together, the results in the chapter suggest that the different *perA* alleles described may encode proteins with higher product diversity than was historically anticipated. Evidence supporting synthesis of a novel product by PerA-ΔR proteins is supplied, and the variability in substrate-binding residues and recombination patterns in PerA-ΔR proteins suggests it is possible that more than one unknown product awaits discovery.

5 Identification and characterisation of the symbiotically regulated gene *irlA*

5.1 Introduction: transcriptome analyses identifies a gene of interest

A previous study by Eaton et al. (2010) compared transcriptomes of *L. perenne* plants infected with either wt *E. festucae* Fl1 or the symbiosis-defective Δ*sakA* mutant (Eaton et al., 2008). The aim of this study was to identify differentially regulated genes between the two associations that represent candidates for symbiotically important genes. One finding from this study was the identification of a symbiotically regulated gene cluster putatively involved in nitrogen metabolism. A subsequent study by Eaton et al. (2015) compared the transcriptomes of *L. perenne* plants infected with either wt or one of the Δ*sakA*, Δ*proA* or Δ*noxA* mutants. Similarly to the Δ*sakA* mutant, the Δ*proA* and Δ*noxA* mutants had previously been shown to exhibit strong symbiosis-defective phenotypes *in planta* (Tanaka et al., 2013; Tanaka et al., 2006). One of the putative nitrogen metabolism cluster genes from the original transcriptome analysis (Eaton et al., 2010), encoding a putative NmrA-like protein, was identified as one of only 39 genes across the entire *E. festucae* Fl1 genome for which expression was significantly downregulated in all three symbiosis-defective mutants (3.9-fold, 38-fold and 5.9-fold downregulation relative to wt in the Δ*sakA*, Δ*proA* and Δ*noxA* mutants, respectively). This gene was also significantly downregulated in the transcriptome of *L. perenne* plants infected with the symbiosis-defective *E. festucae* Fl1 Δ*hepA* mutant relative to wt (6.5-fold downregulation; Chujo et al., unpublished data). Additionally, comparison of RNA-seq data generated from wt grown in culture to wt grown *in planta* found expression of this gene was dramatically higher *in planta* (162-fold upregulation; Chujo et al., unpublished data). Together, these analyses strongly suggested a core role in the *E. festucae*-*L. perenne* symbiosis for this *nmrA*-like gene.

5.2 Results: Bioinformatic analysis of a putative *nmrA*-like gene

NmrA is an *Aspergillus nidulans* protein that mediates post-translational repression of the alternative nitrogen metabolism pathway-activating transcription factor AreA (Stammers et al., 2001). The symbiotically regulated *E. festucae* Fl1

nmrA-like encoding gene was not predicted to contain introns, and this was confirmed by analysis of RNA-seq data from Fl1-infected *L. perenne* plants. Searches of the GenBank database supported annotation of this *E. festucae* Fl1 gene as encoding a 319 aa NmrA-like protein, with the majority of the closest 100 BLASTp sequence matches annotated as hypothetical (60% of matches) or as belonging to the NmrA-like family (26% of matches). However, 8% of the closest 100 sequences identified using BLASTp were annotated as isoflavone reductase-like (IRL). As there were no significant BLASTp matches to previously characterised proteins, the accuracy of these annotations was uncertain. BLASTp searches also indicated that sequences related to this *E. festucae* NmrA-like query were widespread throughout the filamentous fungi.

5.2.1 *In silico* analysis defines *irlA*, a gene encoding an isolflavone reductase-like protein

To evaluate the relative merits of annotating this *E. festucae* gene as either "*nmrA*-like" or "*isoflavone reductase*-like" (IRL), the encoded protein sequence was aligned (2.13.1) with NmrA from *Aspergillus nidulans* (Andrianopoulos et al., 1998), the *Neurospora crassa* NmrA homologue Nmr-1 (Fu et al., 1988), isoflavone reductase (IFR) from *Medicago sativa* (common name: alfalfa; Wang et al., 2006) and an IRL protein from *Arabidopsis thaliana* (Wang et al., 2006). Phylogenetic reconstruction of these sequences (2.13.2) revealed that this *E. festucae* protein was more closely related to the IFR and IRL proteins from plants than the NmrA-family proteins from filamentous fungi (Figure 5.1). The implication that this gene encodes an IRL protein was further supported by the conservation of both an NADPH binding motif and a catalytic lysine residue (Appendix 9.6.1), which are essential for protein function in *M. sativa* IFR (Wang et al., 2006). These residues are also essential for function in PIP family proteins, within which the IRL subfamily is located, such as eugenol synthase (EGS; Louie et al., 2007), pinoresinol-lariciresinol reductase (PLR; Min et al., 2003) and phenylcoumaran benzylic ether reductase (PCBER; Min et al., 2003), but are not conserved in NmrA-like proteins (Andrianopoulos et al., 1998; Stammers et al., 2001). The protein sequence encoded by this *E. festucae* gene was also analysed using the Phyre2 protein-modelling server (Kelley et al., 2015), which

identified EGS and IFR among the proteins with the highest similarity to the predicted structure of this protein query. This *E. festucae* gene was assigned the provisional name *irlA*, encoding the putative isoflavone reductase-like protein IrlA.

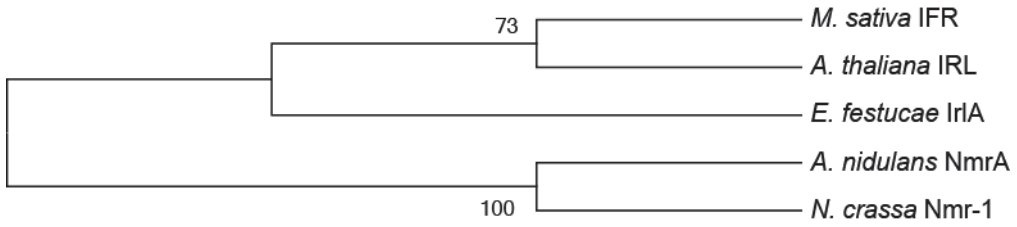


Figure 5.1 Plylogeny analysis places IrlA with plant IFR proteins. Bootstrap consensus tree for primary sequences of *Medicago sativa* isoflavone reductase (IFR), an *Arabidopsis thaliana* isoflavone reductase-like protein (IRL), the *Epichloë festucae* IrlA protein, the *Aspergillus nidulans* NmrA protein and the *Neurospora crassa* NmrA homologue Nmr-1. Phylogenetic tree was generated from a multiple sequence alignment (9.6.1) as described in Section 2.13.2.

5.3 Results: generation and phenotypic analysis of *irlA* deletion mutants

Targeted gene replacement by homologous recombination was utilised to investigate the role played by the IrlA protein in the *E. festucae-L. perenne* symbiosis. Mutants were generated in which a P*trpC-hph* expression cassette replaced the entire *irlA* coding sequence, with *hph*-derived resistance against the antibiotic hygromycin providing a selectable marker for transformants.

5.3.1 Generation of transformants 1-200 from a pDB02-derived *irlA* deletion construct

The *irlA* deletion construct (Figure 5.2A) was excised by restriction enzyme digestion (2.6.1) from pDB02 (9.3.5) using *Eco*RI, then purified by gel extraction (2.6.9). The purified deletion construct was used to transform *E. festucae* Fl1 protoplasts (2.9.2), and 200 colonies growing through the hygromycin-containing selective overlay were picked for further analysis. Each transformant was sub-

cultured three times to confirm transformant stability and to promote nuclear purity. Crude gDNA was extracted from each transformant (2.5.5) and analysed by a PCR screen (2.6.3) using primer pair irlA_scr_3/irlA_scr_4. High-quality gDNA was then extracted from those transformants identified as Δ*irlA* candidates. This gDNA was used as template for PCR amplification from within the *hph* expression cassette across the 5' and 3' homologous flanks of the deletion construct using primer pairs irlA_scr_1/irlA_scr_2 and irlA_scr_5/irlA_scr_6, respectively. These PCR amplifications confirmed transformants 62, 64, 133 and 172, as Δ*irlA* mutants. High quality gDNA from transformants 62, 64, 133 and 172 was then digested with *Pvu*I or *Pvu*II to generate Southern blots (2.8.1). These blots were probed (2.8.3) with DIG-dUTP-labelled (2.8.2) *irlA* deletion construct linearised by *Eco*RI digestion. The hybridised blots were developed (2.8.4), revealing that mutant Δ*irlA*-62 contained a 2-copy concatamer of the *irlA* deletion construct integrated at the target locus, while Δ*irlA*-64, Δ*irlA*-133 and Δ*irlA*-172 contained n > 2-copy concatamers integrated at the target locus (Figure 5.2). The *Pvu*I Southern also revealed that Δ*irlA*-64 contained an ectopic integration of the deletion construct in addition to the target site integration (Figure 5.2).

5.3.2 Generation of transformants 201-600 from a pDB36-derived *irlA* deletion construct

No single-copy integration Δ*irlA* mutants were obtained from transformants 1-200 (5.3.1). This may have been due to the presence of compatible sticky-ends on the linear DNA fragments used for the transformation causing increased tandem integration frequency by promoting deletion construct concatamerisation. Sequencing across the junction between the two *irlA* deletion constructs in Δ*irlA*-62 supported this hypothesis, with a single fully intact *Eco*RI recognition site identified separating the two constructs (TTAGAATTCCAA; flanking *E. festucae* sequence shown in black). This result indicated that the compatible *Eco*RI-generated sticky ends had ligated. The plasmid pDB36 was therefore generated with the *irlA* deletion construct flanked by an *Eco*RI recognition site at the 5' end and a *Bcl*I recognition site at the 3' end (9.3.22). While it was originally intended that *Bcl*I would be used in conjunction with *Eco*RI to excise the *irlA* deletion construct from pDB36, problems were

encountered obtaining sufficient high quality plasmid DNA from a *dam*$^{-}$ *E. coli* strain, and *Bcl*I cannot digest Dam-methylated DNA. A pre-existing *Bsp*HI recognition site within the *irlA* deletion construct right border was therefore selected to replace *Bcl*I for releasing the 3' end of the *irlA* deletion construct from pDB36. This reduced the size of the right border from 1058 to 901 bp. An additional 400 transformants were generated with this pDB36-derived *irlA* deletion construct using the same transformation and screening strategies described for transformants 1-200 (5.3.1).

Of these new transformants strains 219, 438, 487, 552, 573 and 578 were confirmed as Δ*irlA* mutants by PCR. A Southern blot was then generated for these transformants using *Pvu*I- and *Pvu*II-digested gDNA. This blot was probed with full circular DIG-dUTP-labelled pDB36 plasmid. The results showed that Δ*irlA*-487 was the only strain to contain a single-copy integration of the pDB36-derived *irlA* deletion construct at the target locus (Figure 5.3). Δ*irlA*-487 also did not contain any additional ectopic integrations of the *irlA* deletion construct. Strains Δ*irlA*-438 and Δ*irlA*-573 contained two-copy *irlA* deletion construct concatamers at the target locus with no ectopic integrations, while strains Δ*irlA*-219 and Δ*irlA*-552 exhibited more complex deletions, with several expected bands missing and multiple additional bands indicating ectopic integrations of the *irlA* deletion construct (Figure 5.3). The *Pvu*I digest of strain Δ*irlA*-578 indicated this strain was a tandem integration mutant with no ectopic integrations; however, the *Pvu*II digest showed two bands where only one was expected (Figure 5.3). The smaller of these two bands was consistent with a two-copy tandem insertion, while the larger was consistent with a three-copy tandem insertion, indicating that this strain was a mixed culture of two independent Δ*irlA* mutants.

The identification of six new Δ*irlA* mutants from 400 transformants indicated the 157 bp reduction in right border size did not adversely affect the targeted integration frequency. It was unclear how much of an impact the incompatible sticky ends of this deletion construct had on tandem integration frequency, as while a single-copy integration Δ*irlA* mutant was obtained via this strategy, five additional tandem integration Δ*irlA* mutants were also generated. Sequencing across the junction between the two *irlA* deletion constructs in Δ*irlA*-438 revealed a hybrid of the *Eco*RI/*Bsp*HI recognition sequences (TCATCATGCCAA; *Eco*RI recognition site =

G|AATTC, *Bsp*HI recognition site = T|CATGA, flanking *E. festucae* sequence shown in black), indicating that incompatible sticky ends did not prevent concatamer formation during transformation.

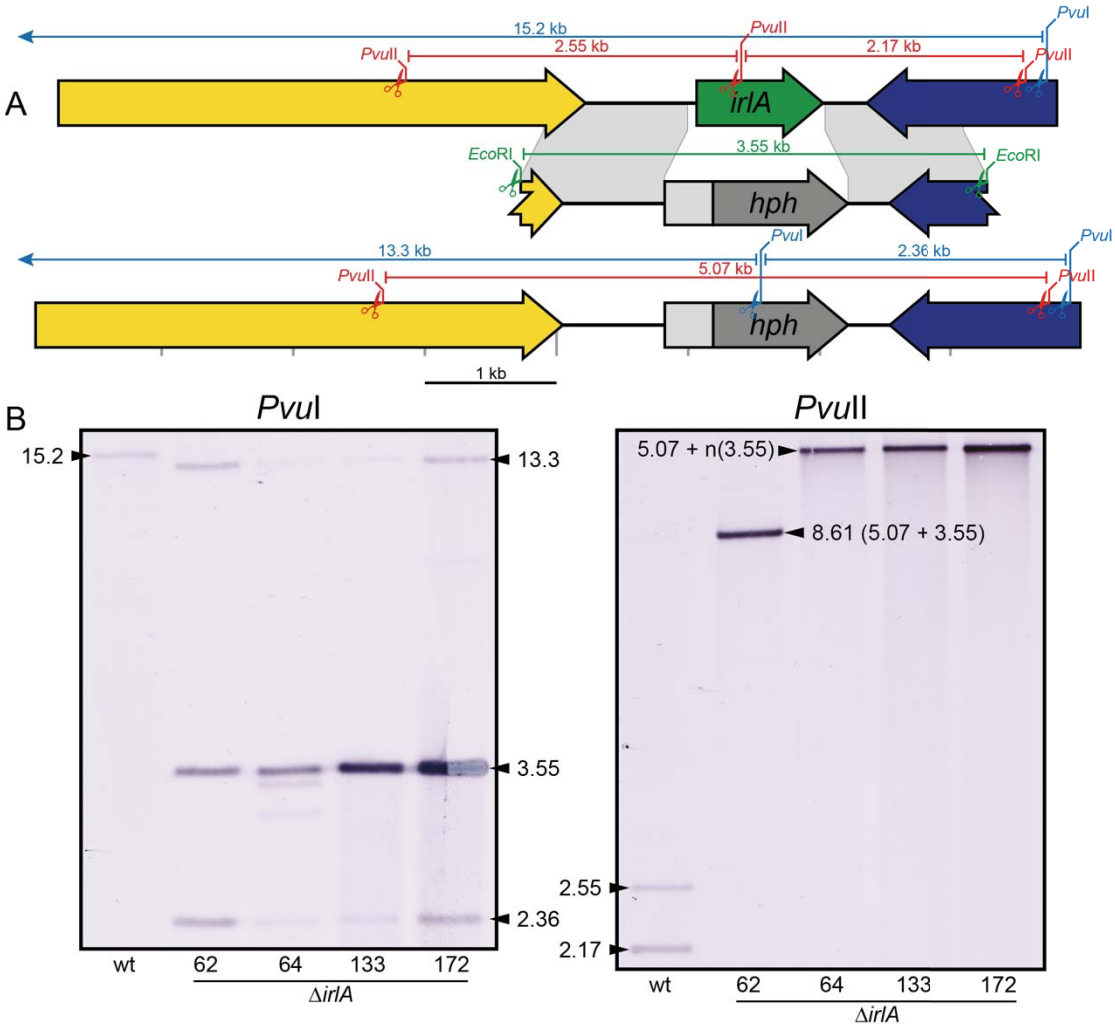


Figure 5.2 Southern blot of pDB02-derived Δ*irlA* mutants. (A) Top: gene map of wt *irlA* locus with fragment sizes expected from digestion with *Pvu*I or *Pvu*II annotated. Middle: gene map of the *irlA* deletion construct post-excision from pDB02 by digestion with *Eco*RI. Bottom: gene map of the anticipated Δ*irlA* locus annotated with 1 kb scale markers and fragment sizes expected post-digestion with *Pvu*I or *Pvu*II. All gene maps are of equal scale. (B) Southern blots of wt, Δ*irlA*-62, Δ*irlA*-64, Δ*irlA*-133 and Δ*irlA*-172 gDNA digested with *Pvu*I (left) or *Pvu*II (right) and probed with DIG-dUTP labelled *irlA* deletion construct excised from pDB02 by *Eco*RI. Band sizes of expected fragments are annotated.

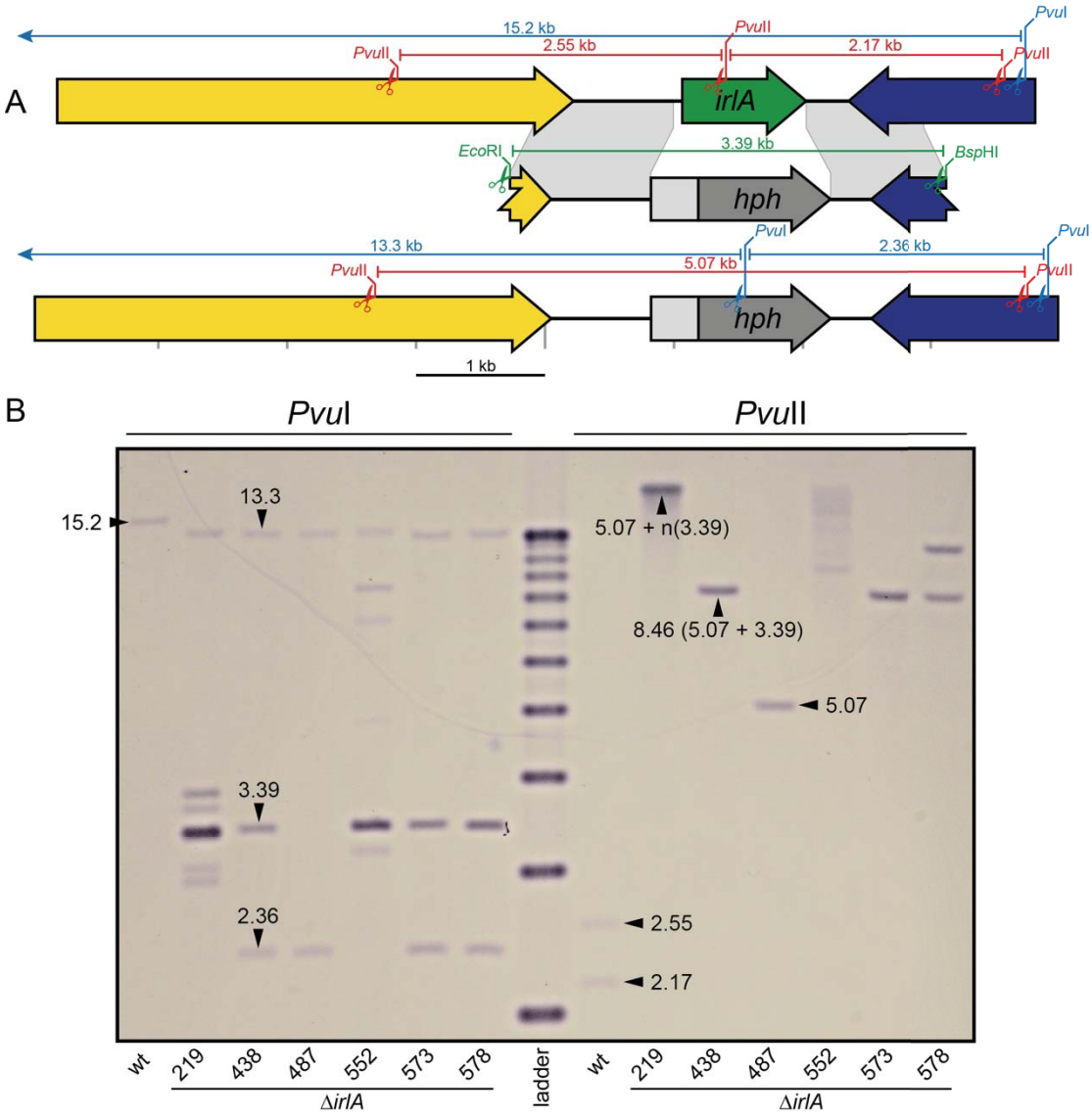


Figure 5.3 Southern blot of pDB36-derived Δ*irlA* mutants. (A) Top: gene map of wt *irlA* locus with fragment sizes expected from digestion with *Pvu*I or *Pvu*II annotated. Middle: gene map of the *irlA* deletion construct post-excision from pDB02 by digestion with *Eco*RI. Bottom: gene map of the desired Δ*irlA* locus annotated with 1 kb scale markers and fragment sizes expected post-digestion with *Pvu*I or *Pvu*II. All gene maps are of equal scale. (B) Southern blot of wt, Δ*irlA*-219, Δ*irlA*-438, Δ*irlA*-487, Δ*irlA*-552, Δ*irlA*-573 and Δ*irlA*-578 gDNA digested with *Pvu*I (left) or *Pvu*II (right) and probed with the DIG-dUTP labelled circular pDB36 plasmid. Band sizes of expected fragments are annotated. The central 1 kb plus ladder (Invitrogen) hybridises to the pDB36-derived probe due to homologous DNA sequences shared between the pDB36 vector backbone and the ladder fragments ≥ 1 kb in size. The smallest (lowest) ladder band in this image is 2 kb in size, and each subsequent band (ascending) is 1 kb larger than the previous band. The largest (highest) band of the 1 kb plus ladder is 12 kb in size.

5.3.3 Analysis of Δ*irlA* mutant phenotypes

The phenotype of Δ*irlA* mutants was assessed both in culture and *in planta* to investigate the role of IrlA in the *E. festucae-L. perenne* symbiosis. The phenotypes of pDB02-derived and pDB36-derived Δ*irlA* mutants are presented separately as the pDB02-derived mutants predate the pDB36-derived mutants by two years, and significant differences in the phenotypes of these two sets were observed.

5.3.4 Analysis of pDB02-derived Δ*irlA* mutant phenotypes

The colony growth rate and morphology of all Δ*irlA* mutants grown on solid media was indistinguishable from wt (Figure 5.4). Plants infected with these Δ*irlA* mutants exhibited host stunting, hyper-tillering and severe root growth reduction (Figure 5.5). Samples from the pseudostem of Δ*irlA*-62 infected plants analysed by confocal laser scanning microscopy (CLSM) revealed hyper-proliferation of hyphae within the host apoplast and apparent unmasking of chitin in the cell wall of endophytic hyphae (Figure 5.6).

As Δ*irlA*-62 had the lowest *irlA* deletion construct copy number of the four pDB02-derived tandem integration mutants it was selected for detailed analysis. This mutant was therefore considered the least likely to suffer from copy number-associated complications. The results show that the average tiller length for plants infected with the Δ*irlA* mutant was significantly smaller than those plants infected with wt (Figure 5.7). The infection frequency of plants that had been inoculated with Δ*irlA*-62 (13%) was also much lower than those inoculated with wt (40%). Pseudostem samples from plants infected with wt and Δ*irlA*-62 endophyte were analysed by CLSM, again revealing hyper-proliferation of Δ*irlA*-62 hyphae within the host apoplast (Figure 5.8). Unlike the results presented in Figure 5.6, chitin unmasking was not observed in these images, perhaps because the plants from these two experiments were grown for different lengths of time post-inoculation before sampling (Figure 5.8). As the hyphae of *E. festucae* symbiotic mutants are often unable to form lateral fusions in culture (Becker et al., 2015; Kayano et al., 2013), this phenotype was also tested in Δ*irlA*-62. The results show qualitatively that Δ*irlA*-62 retains the ability to fuse (Figure 5.9).

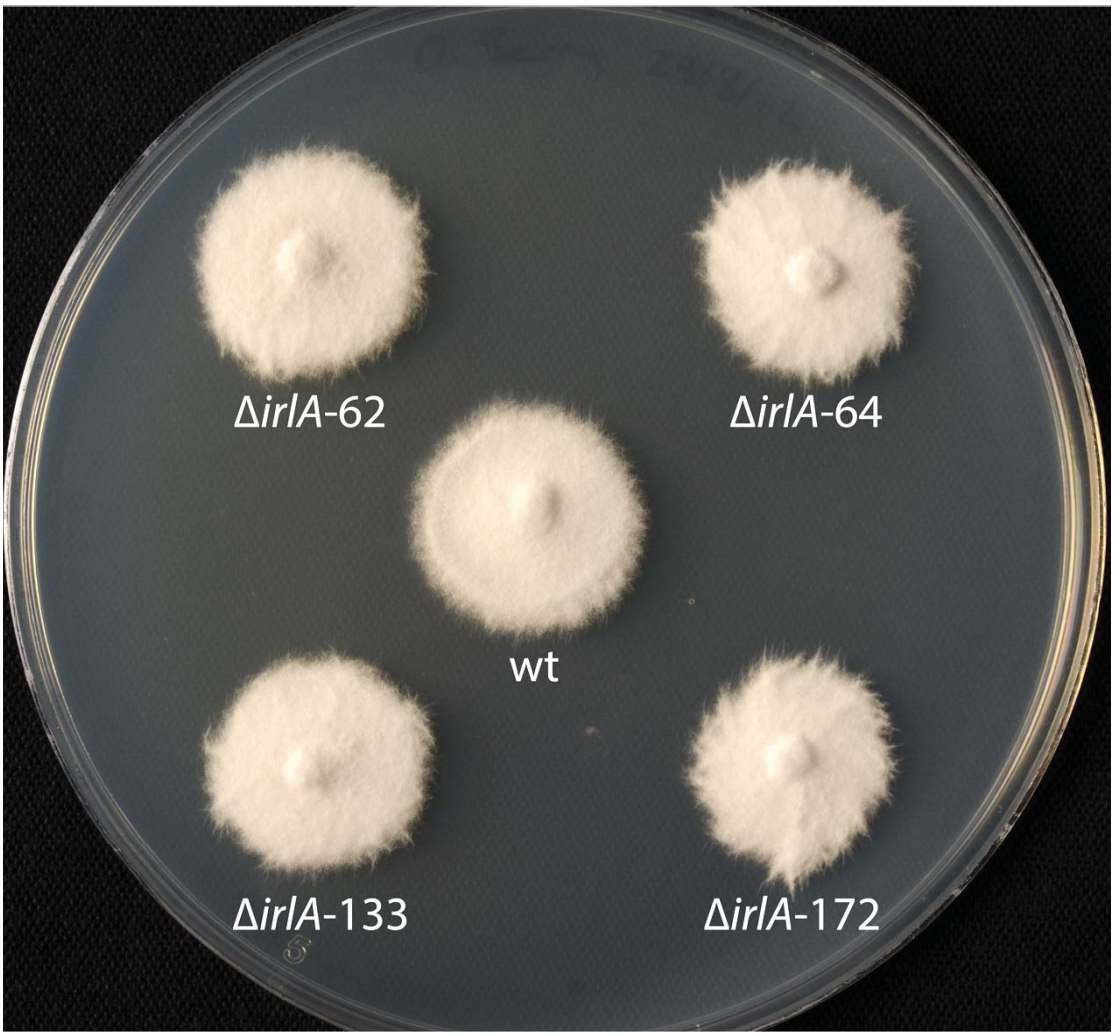


Figure 5.4 Colony morphology of Δ*irlA* mutants. 5 mm diameter circular agar blocks were excised from freshly grown colonies of wild type *E. festucae* Fl1, Δ*irlA*-62, Δ*irlA*-64, Δ*irlA*-133 and Δ*irlA*-172 using sterilised 150 mm glass Pasteur pipettes. These plugs were placed onto a petri dish containing 30 mL of PD media (2.3.3) and grown for 10 days under standard conditions (2.4.2.2) before photographing.

Figure 5.5 Phenotype of *L. perenne* plants infected with pDB02-derived Δ*irlA* mutants. Images of plants infected with mutants Δ*irlA*-62, Δ*irlA*-64, Δ*irlA*-133 and Δ*irlA*-172 (first through fourth plant columns from left of image, respectively) or wt (four plants to right of image) grown for 8 weeks post-inoculation (2.10.1) under controlled conditions (2.10.2). Infection status was confirmed by aniline blue staining (2.12.1). Mild, medium and severely stunted examples (top to bottom) are shown for plants infected for each of the Δ*irlA* mutants.

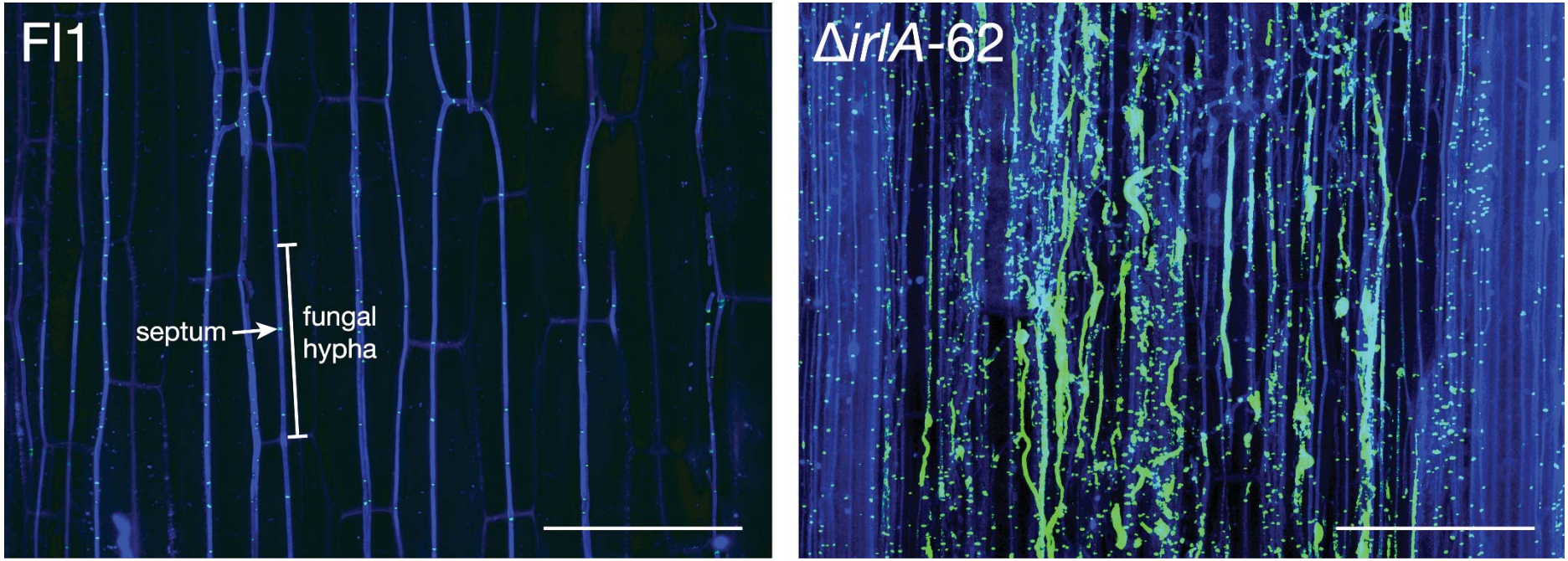


Figure 5.6 Phenotype of Δ*irlA*-62 mutant *in planta*. Confocal laser-scanning microscopy images comparing the phenotypes of wt and Δ*irlA*-62 hyphae growing within *L. perenne* plants at 10 weeks post-inoculation. Hyphae were stained (2.12.3) to visualise the fungal cytoplasm (blue) and chitin (green). Image stacks from 10 µm-thick sections just below the pseudostem tissue surface shown (slices taken at 0.1 µm intervals for Fl1, 0.5 µm for Δ*irlA*-62). Scale bar = 100 µm.

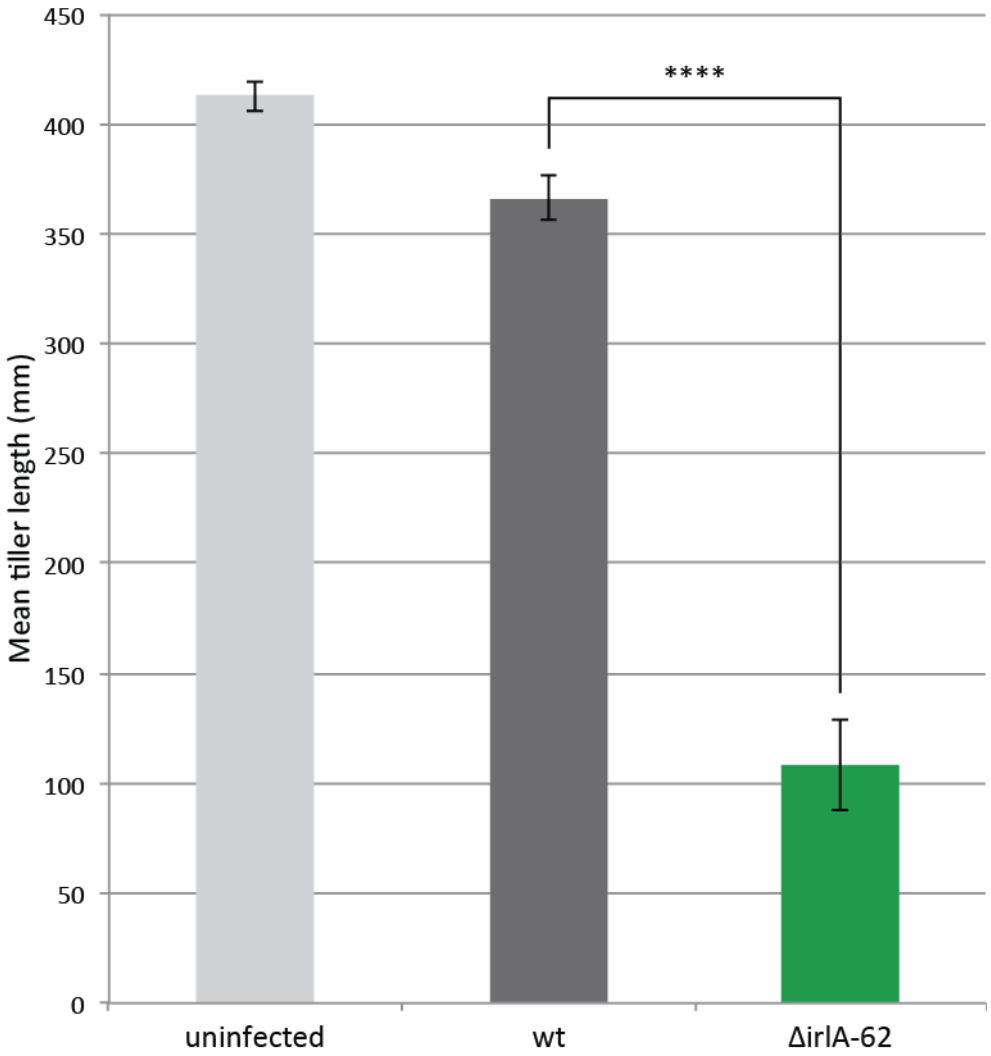


Figure 5.7 Mean tiller length for uninfected, wt infected and Δ*irlA*-62 infected *L. perenne* plants. Graph shows average tiller length from 130 uninfected, 17 wt-infected and 5 Δ*irlA*-62-infected *L. perenne* plants. Uninfected plants statistics were gathered from plants that underwent the initial endophyte inoculation procedure, but were subsequently found to be uninfected by immunoblot analysis (2.10.3). Error bars show the standard error of the mean. The statistical significance of the difference between the wt and Δ*irlA*-62 means was determined using one-tailed Welch's *t*-test ($P < 0.0001$).

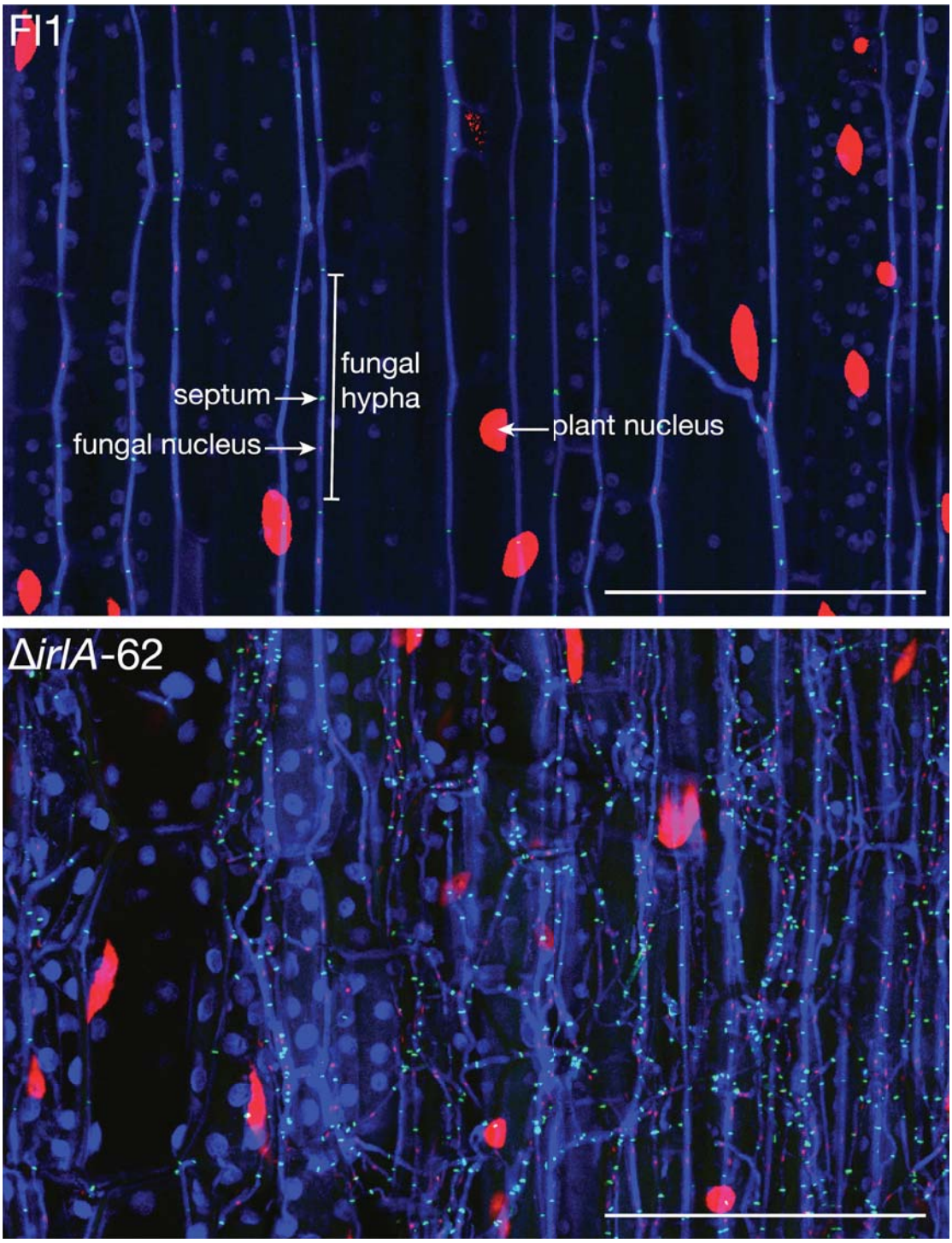


Figure 5.8 Phenotype of Δ*irlA*-62 mutant *in planta*. Confocal laser-scanning microscopy images comparing the phenotypes of wt and Δ*irlA*-62 hyphae growing within *L. perenne* plants at 14 weeks post-inoculation. Samples were stained (2.12.3) to visualise the fungal cytoplasm (blue), chitin (green) and plant nuclei (red). Image stacks from 10 µm-thick sections just below the pseudostem tissue surface shown (slices taken at 0.42 µm intervals for Fl1, 0.38 µm for Δ*irlA*-62). Scale bar = 100 µm.

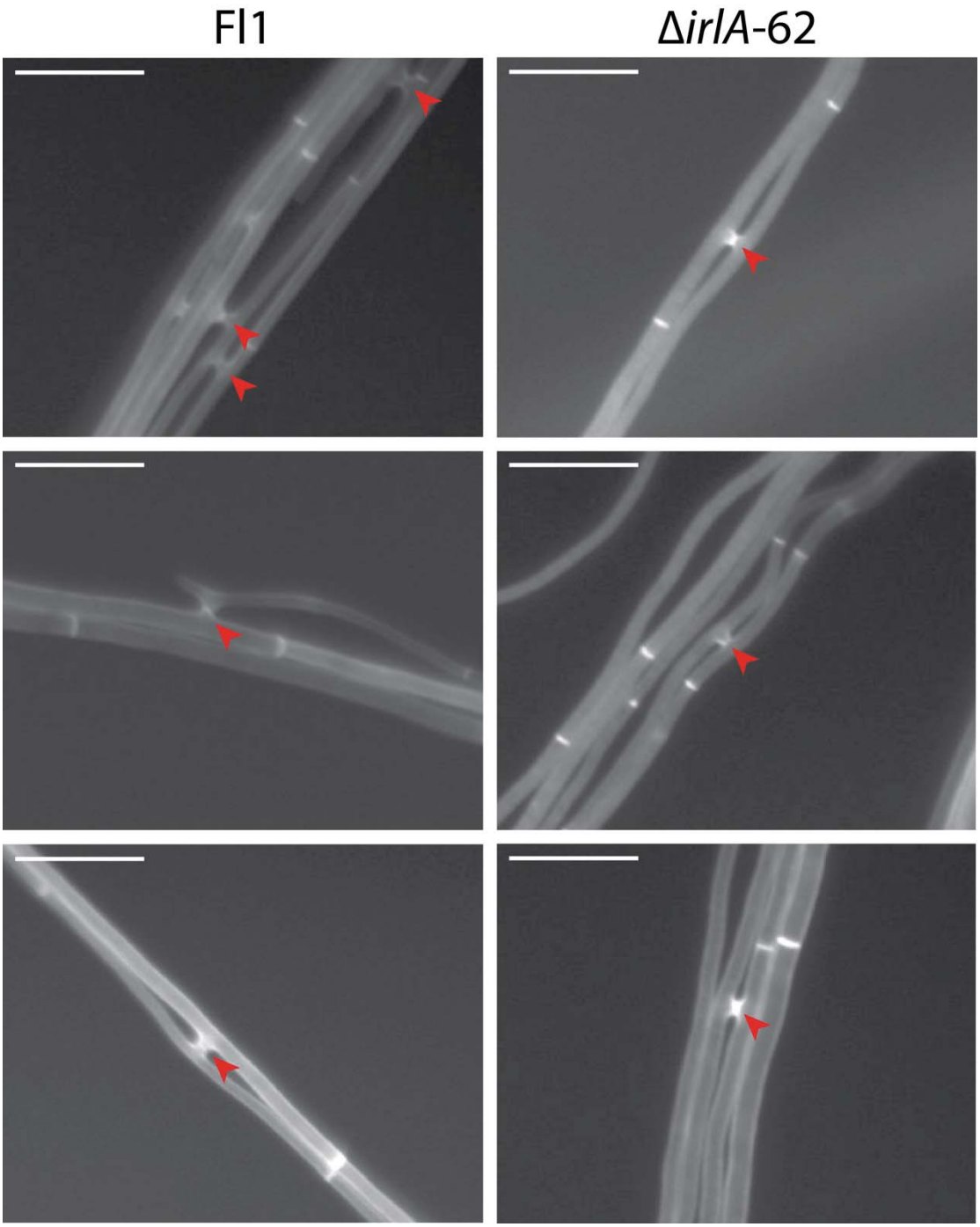


Figure 5.9 Δ*irlA*-62 mutants retain the ability to fuse in culture. Inverted microscope images of wt Fl1 (left) and Δ*irlA*-62 (right) hyphae grown on water agar slides and stained with calcafluor white (2.12.1). Calcafluor stains chitin in the fungal cell wall, and strong staining is observed at fungal septa and inter-hyphal fusion sites. Fusion sites are indicated with red arrowheads. Scale bar = 20 µm.

5.3.5 Ectopic reintroduction of irlA failed to complement the ΔirlA-62 phenotype

Protoplasts generated from strain Δ*irlA*-62 (2.9.1) were transformed with plasmids pDB08 (9.3.7) or pDB09 (9.3.8), which contain the native *irlA* promoter and terminator controlling expression of either *irlA* or an *irlA-eGFP* fusion construct, respectively. As these plasmids do not contain an antibiotic resistance gene suitable for selection in filamentous fungi they were co-transformed with plasmid pII99, which confers resistance to the antibiotic geneticin. These transformants were then screened by extraction of crude gDNA (2.5.5) that was used as template for PCR amplification (2.6.3) with primer pair pDB08_1/pDB08_2. This primer pair was designed for PCR amplification of the entire *irlA* complementation construct. Successful PCR amplification therefore indicated that the genome of the corresponding transformant contained a fully intact *irlA* or *irlA-eGFP* expression construct that should complement the Δ*irlA* genotype. Three PCR-positive pDB08 transformants (strains 08-1, 08-2 and 08-3) and three PCR-positive pDB09 transformants (strains 09-1, 09-2 and 09-3) were selected, inoculated into *L. perenne* seedlings (2.10.1) and the seedlings were grown for 9 weeks post-inoculation under standard conditions (2.10.2). The total number or tillers and length of each tiller were recorded for each plant, followed by immunoblotting to determine infection status (2.10.3). The results showed that plants infected with both the unmodified Δ*irlA*-62 mutants and Δ*irlA*-62-derived pDB08 and pDB09 transformants exhibited consistently elevated tiller number, reduced mean tiller length and reduced infection frequency relative to wt-infected plants (Table 5.1). This indicated that the ectopic reintroductions of *irlA* or *irlA-eGFP* expression constructs in strains 08-1, 08-2, 08-3, 09-1, 09-2 and 09-3 were insufficient to complement the *in planta* Δ*irlA*-62 phenotype in those transformants. Due to the added complication of the tandem deletion construct insertion in Δ*irlA*-62 it was decided to do any future complementation experiments in a single-copy integration Δ*irlA* strain.

5.3.6 Analysis of pDB36-derived Δ*irlA* mutant phenotypes

As the pDB02-derived Δ*irlA* mutants all contained concatamers of the deletion construct at the target locus, a second round of transformation was done with a pDB36-derived deletion construct to obtain a single-copy integration mutant (5.3). Of the six additional Δ*irlA* mutant strains identified from this second transformation round, only Δ*irlA*-487 contained a single copy integration of the *irlA* deletion construct at the target locus (Figure 5.3).

Table 5.1 Host interaction phenotype of Δ*irlA*-62 complementation strains

Endophyte strain	# Plants alive	# Plants infected	Infection frequency	Mean tiller length (mm)	Mean # tillers
wt	15	6	0.40	273	2.3
Δ*irlA*-62	16	4	0.25	112	5.8
08-1	11	1	0.09	130	4.0
08-2	12	2	0.17	119	5.5
08-3	13	1	0.08	281	7.0
09-1	12	0	0.00	NA	NA
09-2	11	1	0.09	108	5.0
09-3	14	2	0.14	150	4.5
uninfected	136	NA	NA	355	3.0

5.3.6.1 Analysis of single-copy integration mutant Δ*irlA*-487

Comparison of plants infected with wt or the single-copy integration mutant Δ*irlA*-487 did not identify any major macroscopic differences (Figure 5.10). CLSM analysis also revealed that growth of Δ*irlA*-487 hyphae *in planta* is similar to wt (Figure 5.11). This asymptomatic Δ*irlA*-487 host-association phenotype and normal infection rate (52% for Δ*irlA*-487, 40% for wt) was unexpected, as it sharply contrasted the symbiosis-defective phenotype and low infection rate previously observed for plants infected with four independent pDB02-derived Δ*irlA* mutants (5.3.4). Several additional experiments were therefore performed to confirm the validity of this result. Firstly, the deletion construct contained within pDB36 was comprehensively resequenced (2.6.13), confirming that there were no SNPs or indels contained within the sequences flanking the P*trpC-hph* cassette that could

potentially disrupt the adjacent genes. This was important, as inactivation of the adjacent genes has the potential to complement the Δ*irlA* phenotype *in planta* (see Chapter 6 for details). Next, the Δ*irlA* locus from Δ*irlA*-487 was amplified by PCR (2.6.3) using primer pair irlA_locus_F/irlA_locus_R and sequenced (2.6.13), with sequence analysis confirming that integration of the deletion construct at the target locus did not induce any unanticipated changes in the flanking genes. Finally, the endophyte from three Δ*irlA*-487 infected plants was re-isolated from plant material (2.10.4) and allowed to grow onto PD media (2.3.3) lacking hygromycin. These isolates were then sub-cultured onto PD media (2.3.3) containing 150 µg mL^{-1} hygromycin, on which all isolates grew normally. This reconfirmed the identity of these isolates as recombinant in origin (ie: Δ*irlA*-487), excluding the possibility that contamination of nominally endophyte-free seed with non-GMO background endophyte was confounding these results.

5.3.6.2 Analysis of pDB36-derived tandem integration Δ*irlA* mutants

As the single-copy insertion mutant Δ*irlA*-487 did not exhibit the expected host-association phenotype, additional Δ*irlA* mutants generated from the shortened pDB36-derived *irlA* deletion construct (5.3.2) were analysed to determine if the host stunting phenotype was specific to tandem integration Δ*irlA* mutants in general, or pDB02-derived tandem integration mutants specifically. The results showed that these pDB36-derived tandem integration Δ*irlA* mutants did not stunt the host plant (Figure 5.12). Immunoblot analysis (2.10.3) of endophyte-inoculated plants determined infection rates of 37% for Δ*irlA*-438-inoculated plants, 36% for Δ*irlA*-573-inoculated plants and 40% for wt-inoculated plants. This showed that the symbiosis-defective phenotypes and lowered infection rates observed for pDB02-derived tandem integration Δ*irlA* mutants were not observed for pDB36-derived tandem integration Δ*irlA* mutants.

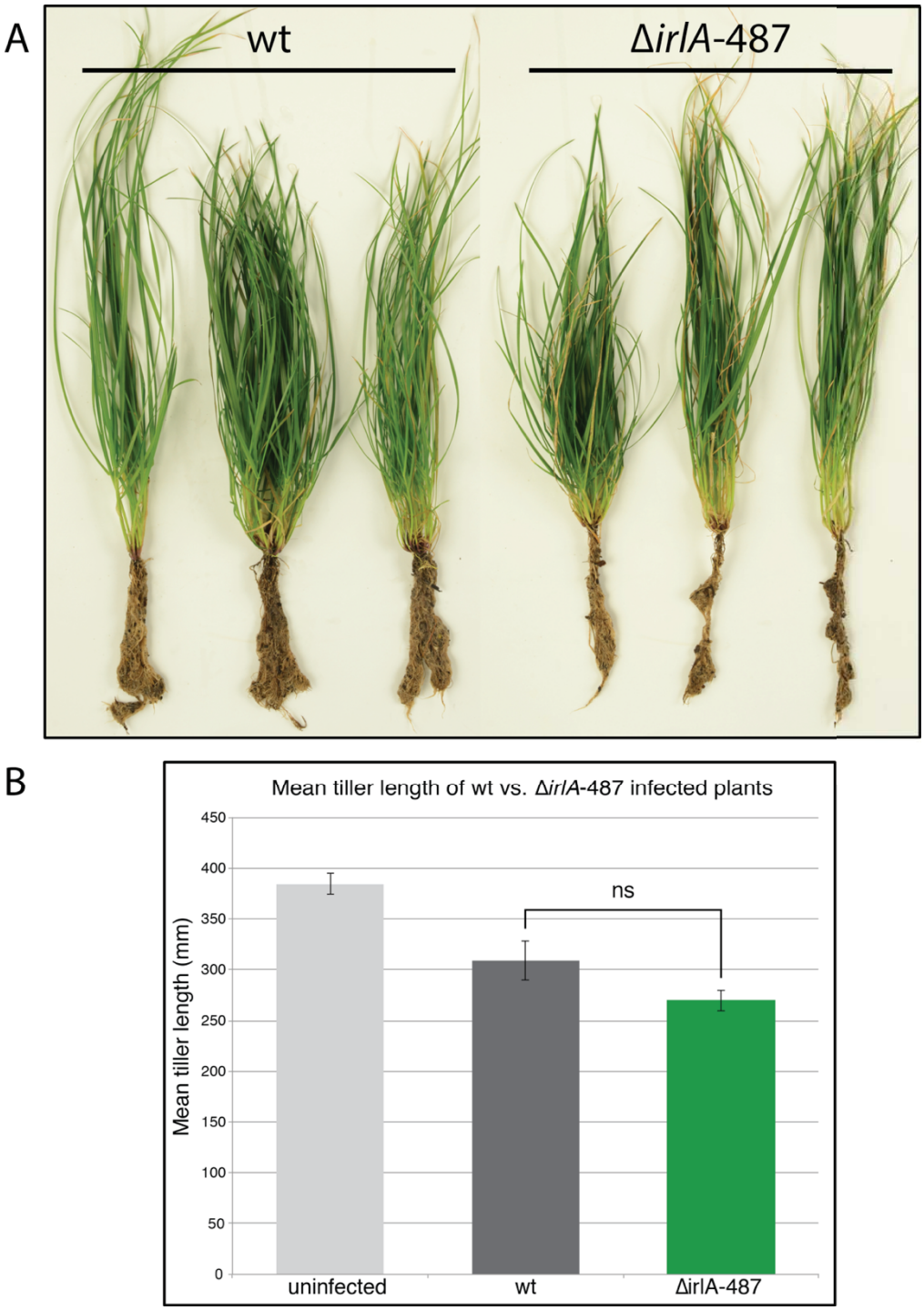


Figure 5.10 Phenotype of Δ*irlA*-487 infected *L. perenne* plants. Δ*irlA*-487 and wt were inoculated into 56 and 28 *Lolium perenne* seedlings each (2.10.1), respectively. Plants were grown under standard conditions (2.10.2) for 11-weeks post inoculation, following which the length of every tiller on each surviving plant was recorded and infection status was determined by immunoblot (2.10.3). (A) Representative wt and Δ*irlA*-487 infected *L. perenne* plants at 12 weeks post-inoculation. (B) Graph showing mean tiller length of 51 uninfected, 7 wt-infected and 23 Δ*irlA*-487-infected *L. perenne* plants. Error bars show the standard error of the mean. The difference in mean tiller length between wt and Δ*irlA*-487 is not significant (ns; one-tailed Welch's *t*-test).

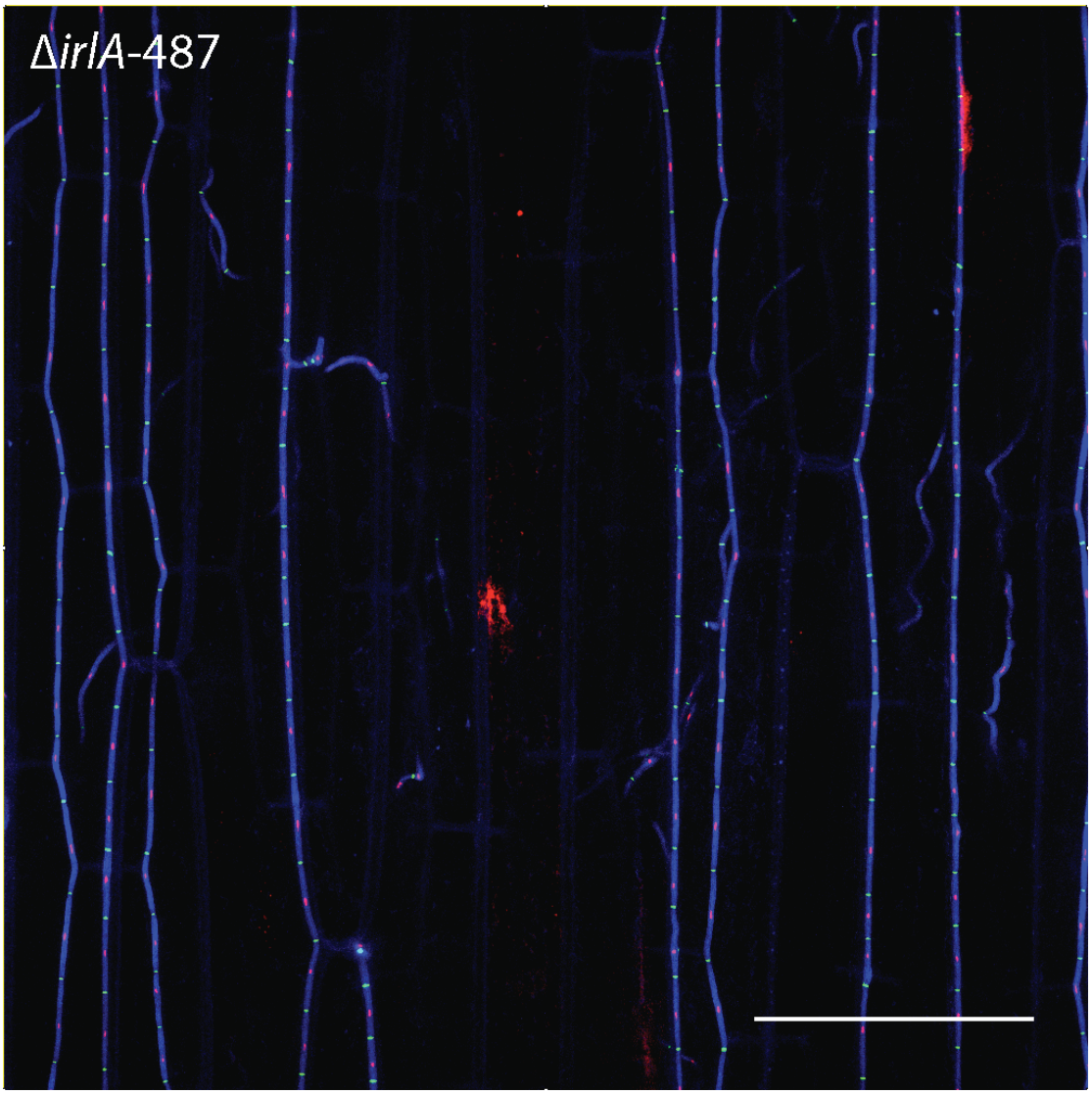


Figure 5.11 Phenotype of Δ*irlA*-487 mutant *in planta*. CLSM images of Δ*irlA*-487 hyphae growing within *L. perenne* plants at 12 weeks post-inoculation. Samples were stained (2.12.3) to visualise the fungal cytoplasm (blue), chitin (green) and nuclei (red). Image stack taken from 10 µm-thick section just below the pseudostem tissue surface (slices taken at 0.5 µm intervals). Scale bar = 100 µm.

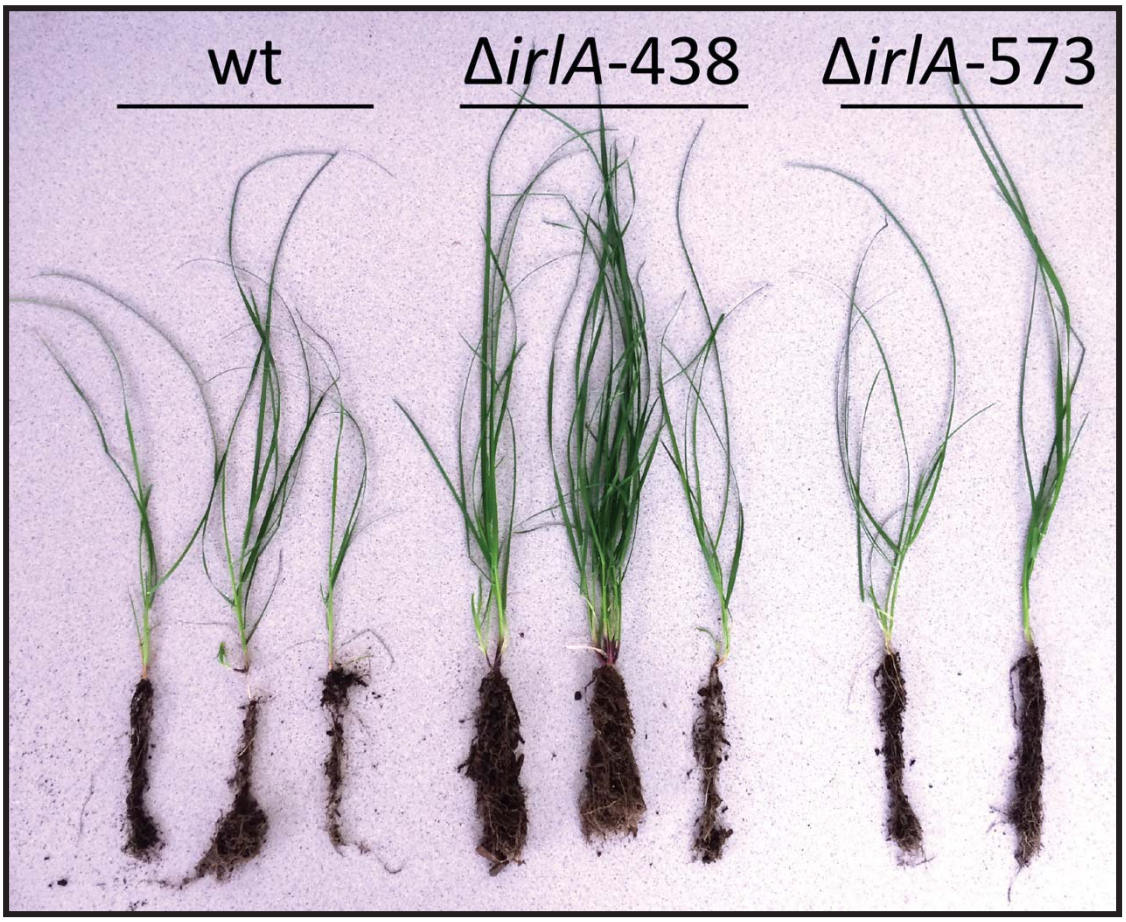


Figure 5.12 Phenotype of *L. perenne* plants infected with pDB36-derived tandem integration Δ*irlA* mutants. Wt, Δ*irlA*-438 and Δ*irlA*-573 were inoculated (2.10.1) into 28 seedlings each and grown under standard conditions (2.10.2). Infection status was then confirmed by immunoblot (2.10.3). Photo shows representative examples of wt, Δ*irlA*-438 and Δ*irlA*-573 infected plants at 9 weeks post-inoculation.

5.4 Results: comparing *irlA* locus gene expression between pDB02-derived and pDB36-derived Δ*irlA* mutants

In Chapter 6 it is explained that *irlA* is part of a proposed five-gene cluster (Figure 5.13), and that increasing the expression of other genes in this cluster during vegetative growth *in planta* might be expected to cause a symbiosis-defective phenotype. It was therefore decided to investigate the expression profile of *irlA* and the other genes in this cluster (*pdtA*, *afrA*, *mfsB* and *fxbA*) across Δ*irlA*-62, Δ*irlA*-133, Δ*irlA*-438, Δ*irlA*-487, Δ*irlA*-573 and wt to determine if differential expression of these genes could be responsible for the *in planta* phenotypic disparities between pDB02-derived and pDB36-derived Δ*irlA* mutants.

5.4.1 Analysis of *irlA* locus gene expression in liquid culture

Liquid cultures of mutants Δ*irlA*-62, Δ*irlA*-133, Δ*irlA*-438, Δ*irlA*-487, Δ*irlA*-573 and wt were grown in triplicate under standard conditions (2.4.2.1) for 5 days, following which mycelia were harvested (2.4.4) and snap frozen in liquid nitrogen. Approximately 100 mg from each sample was used for RNA extraction (2.7.1), and 600 ng of extracted RNA from each sample was used to generate cDNA (2.7.3). 5-fold dilutions of these cDNA samples were then used as template for 40-cycle qPCR reactions (2.6.6) using the primer pairs described in Table 5.2, for which target sequences are visualised in Figure 5.13. Analysis of RNAseq data had shown that expression of all *irlA* cluster genes was extremely low in culture (T. Chujo, unpublished data). The RT-qPCR results from all samples showed that the target sequences of primer pairs 1 (*pdtA*), 2 (*afrA*) and 7 (*fxbA*) were either not detected or detected at lower levels than the least concentrated standard, and were not significantly upregulated in any of the Δ*irlA* mutants when compared to wt (Figure 5.14). In contrast, the target sequence of primer pair 3 (located 3' of primer pair 2 within *afrA*) was significantly upregulated in all Δ*irlA* mutants except Δ*irlA*-487 (Figure 5.14). The target sequence of primer pair 5 (*mfsB*) was significantly upregulated in all Δ*irlA* mutants, though to a lesser extent in Δ*irlA*-487 (Figure 5.14). The target sequence of primer pair 6 (located 5' of primer pair 5 within *mfsB*) was significantly upregulated in pDB36-derived Δ*irlA* mutants only, though expression

levels were considerably reduced for mutants Δ*irlA*-438 and Δ*irlA*-573 relative to the primer pair 5 target (Figure 5.14). This tandem deletion mutant-specific disparity in expression levels across different regions of both *afrA* (target sequences of primer pairs 2 and 3) and *mfsB* (target sequences of primer pairs 5 and 6) was proposed to be due to incomplete termination of P*trpC*-driven *hph* expression. This would cause read-through into the *afrA* and *mfsB* gene fragments that flank each junction between individual units of the deletion construct concatamer contained within each tandem integration Δ*irlA* mutant. This hypothesis was supported by the positive correlation between *irlA* deletion construct copy number and expression level of the primer pair 3 and 5 target sequences, which are located within the left and right borders of all Δ*irlA* deletion constructs, respectively (Figure 5.13, Figure 5.14). This positive relationship between copy number and gene expression was also observed for *hph* (9.4.3). This hypothesis was confirmed by detection of significant expression levels for an *afrA*-*mfsB* RNA hybrid in both pDB02-derived and pDB36-derived tandem integration Δ*irlA* mutants using primer pairs 9 and 10, respectively (Figure 5.14). Additionally, P*trpC*-driven read-through into the downstream *mfsB* gene in the antisense direction explains why the expression levels of the primer pair 5 target are elevated in the single-copy Δ*irlA*-487 mutant relative to wt (Figure 5.14). A similar mechanism would be expected to elevate expression of the primer pair 6 target to similar levels in all Δ*irlA* mutants; however, only the pDB36-derived tandem integration mutants exhibited this increase (Figure 5.14). This might have been due to the antisense *mfsB* RNA binding to sense *mfsB* mRNA in the pDB02-derived tandem integration Δ*irlA* mutants, forming double stranded RNA that induced degradation of complementary RNA molecules via the RNA interference pathway (discussed further in Section 5.5). While informative, this gene expression analysis of the *irlA* cluster genes in culture did not identify any obvious differences between pDB02-derived and pDB36-derived tandem integration Δ*irlA* mutants that could explain the host-interaction phenotype disparity between these mutants.

5.4.2 Analysis of *irlA* locus gene expression *in planta*

The expression of *irlA* cluster genes was analysed *in planta* to determine if there were any differences that may explain the disparity in host-interaction

phenotype between pDB02-derived and pDB36-derived Δ*irlA* mutants. *L. perenne* seedlings were inoculated (2.10.1) with wt, Δ*irlA*-62, Δ*irlA*-133, Δ*irlA*-438, Δ*irlA*-487 or Δ*irlA*-573, grown under standard conditions (2.10.2) for 9 weeks post-inoculation then immunoblotted to determine infection status (2.10.3). The plants were allowed to recover for 4 days, following which several 1 cm pseudostem sections were excised from just above the roots of three infected plants for each endopyte strain except Δ*irlA*-62, for which only one infected plant was sampled due to a lower than anticipated infection frequency. These sections were then immediately submersed in liquid nitrogen after removing dead leaf sheaths, and whole RNA was subsequently extracted from these frozen samples (2.7.2). A total of 1 µg RNA from each sample was then used as template for cDNA synthesis (2.7.3), and 1:3 dilutions of this cDNA were used as templates for 40-cycle qPCR reactions (2.6.6) using the primer pairs described in Table 5.2.

The results showed that expression of *pdtA*, *afrA*, and *fxbA* was generally not detectable in wt *E. festucae* Fl1–*L. perenne* associations, while *mfsB* expression was detectable but extremely low (Figure 5.15). In contrast, *pdtA*, *afrA* and *fxbA* expression was significantly increased in associations involving the pDB02-derived Δ*irlA*-62 and Δ*irlA*-133 mutants (Figure 5.15). These genes also generally showed increased expression relative to wt in associations involving the pDB36-derived Δ*irlA*-438 and Δ*irlA*-573 mutants, but at significantly lower levels than associations involving the pDB02-derived mutants (Figure 5.15). Expression of these genes in associations containing the single-copy Δ*irlA*-487 mutant was generally undetectable except for *fxbA*, for which expression was detected at extremely low levels (Figure 5.15). Expression of *mfsB* was higher in associations involving any of the Δ*irlA* mutants relative to wt-infected plants, with the highest levels detected in associations containing the pDB02-derived Δ*irlA*-62 or Δ*irlA*-133 mutants and lowest levels detected in associations containing the single-copy Δ*irlA*-487 mutant (Figure 5.15). Analysis of *mfsB* expression is complicated in these associations by the established occurrence of P*trpC*-driven transcription of antisense *mfsB* RNA (5.4.1), and the inability to differentiate this from sense *mfsB* mRNA. Overall, these results established that expression of the *irlA* cluster genes was substantially higher in associations involving pDB02-derived tandem integration Δ*irlA* mutants compared to

both wt associations and associations containing any of the pDB36-derived tandem integration Δ*irlA* mutants.

Interestingly, the positive correlation observed in culture between *irlA* deletion construct copy number and the concentration of RNA transcripts from sequences within the *irlA* deletion construct was not observed *in planta*. Instead, pDB02-derived mutants showed considerably higher expression of these transcripts *in planta* than pDB36-derived mutants despite Δ*irlA*-62 and Δ*irlA*-438/Δ*irlA*-573 having the same copy number, and this trend also extended to the target of primer pair 6 (Figure 5.15, Appendix 9.4.4). This indicates that P*trpC*-driven transcription is considerably upregulated for pDB02-derived mutants *in planta* relative to pDB36-derived mutants, perhaps due to a change in the heterochromatic state of this region associated with upregulation of the remaining *irlA* cluster genes *in planta*. However, this may also be an artefact caused by the lack of *irlA*-62 replicates.

Table 5.2 Primer pairs used for RT-qPCR analysis of *irlA* locus gene expression in *E. festucae*

ID	Primer 1	Primer 2	Product size (bp)	Target
1	pdtA_3	pdtA_4.1	144	3’ end of *pdtA*
2	afrA_3	afrA_4.1	90	3’ end of *afrA* outside of *irlA* deletion construct left border
3	afrA_5	afrA_6	84	3’ end of *afrA* inside of *irlA* deletion construct left border
4	irlA_3	irlA_4	114	3’ end of *irlA*
5	mfsB_3	mfsB_4.1	110	3’ end of *mfsB* inside of *irlA* deletion construct right border
6	mfsB_8	mfsB_9	113	3’ end of *mfsB* outside of *irlA* deletion construct right border
7	fxbA_3.1	fxbA_4.1	145	3’ end of *fxbA*
8	hph_1	hph_2	132	3’ end of *hph*
9	afrA_7	mfsB_10	136	3’-5’ interface between pDB02-derived *irlA* deletion construct concatamers
10	afrA_7	mfsB_11	133	3’-5’ interface between pDB36-derived *irlA* deletion construct concatamers
11	TC399	TC400	82	3’ UTR of *EF2*
12	TC407	TC408	83	3’ UTR of *rps22*

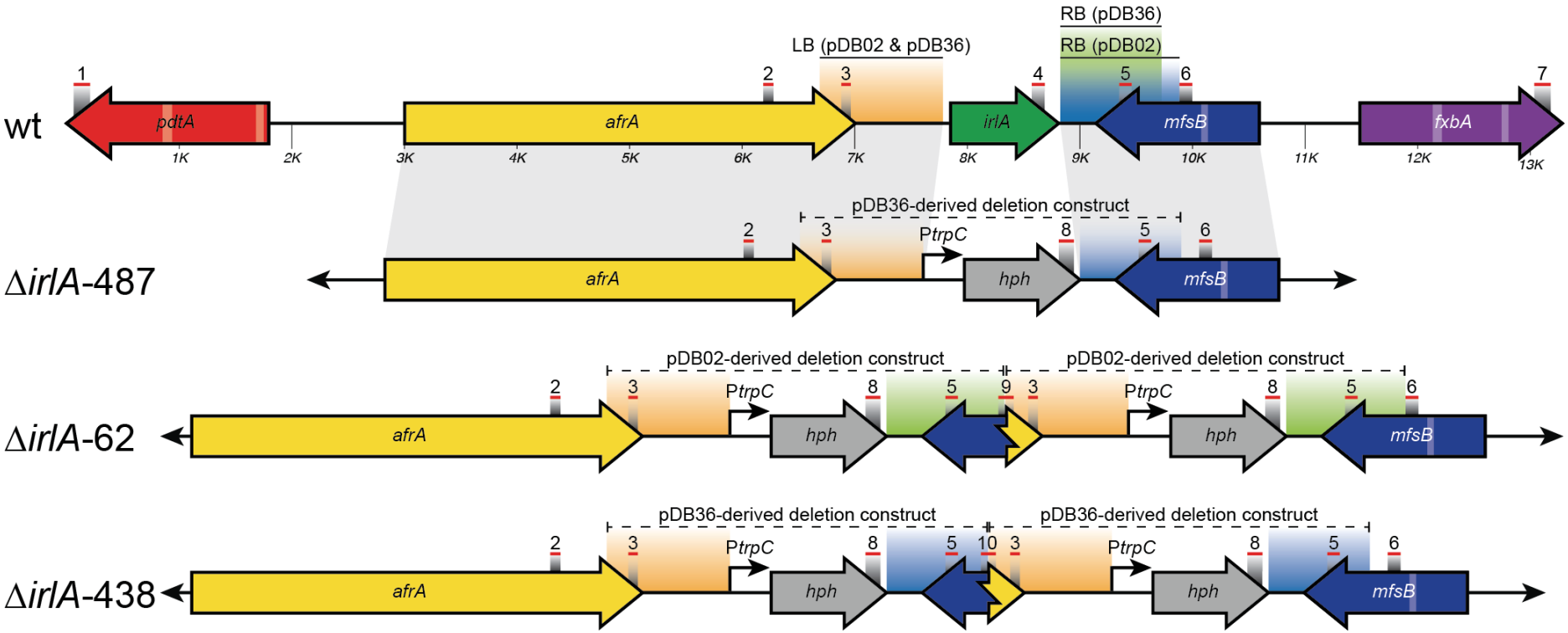


Figure 5.13 Location of regions amplified for RT-qPCR analysis of Δ*irlA* mutants. Same-scale maps of *irlA* locus genes in wt, Δ*irlA*-487, and the two-copy tandem integration mutants Δ*irlA*-62 (derived from pDB02) and Δ*irlA*-438 (derived from pDB36). Gene maps are annotated with numbered red bars indicating regions amplified for gene expression analysis by RT-qPCR (Figure 5.14, Figure 5.15). Also annotated is the region used to generate the left border (LB; orange highlights) and right border (RB; green and blue highlights for pDB02 or pDB36, respectively) of the *irlA* deletion constructs. The 5' and middle sections of both RB fragments are identical, but the 3' end is 157 bp longer in pDB02. Lighter shaded sections of the gene arrows define intron locations. The *Aspergillus niduans trpC* promoter (P*trpC*) is strongly and constitutively expressed in *E. festucae* (Eaton et al., 2010; Eaton et al., 2015). Not shown here are the n-copy pDB02-derived tandem mutant Δ*irlA*-133, and a second pDB36-derived two-copy tandem mutant Δ*irlA*-573 (structurally identical to Δ*irlA*-438).

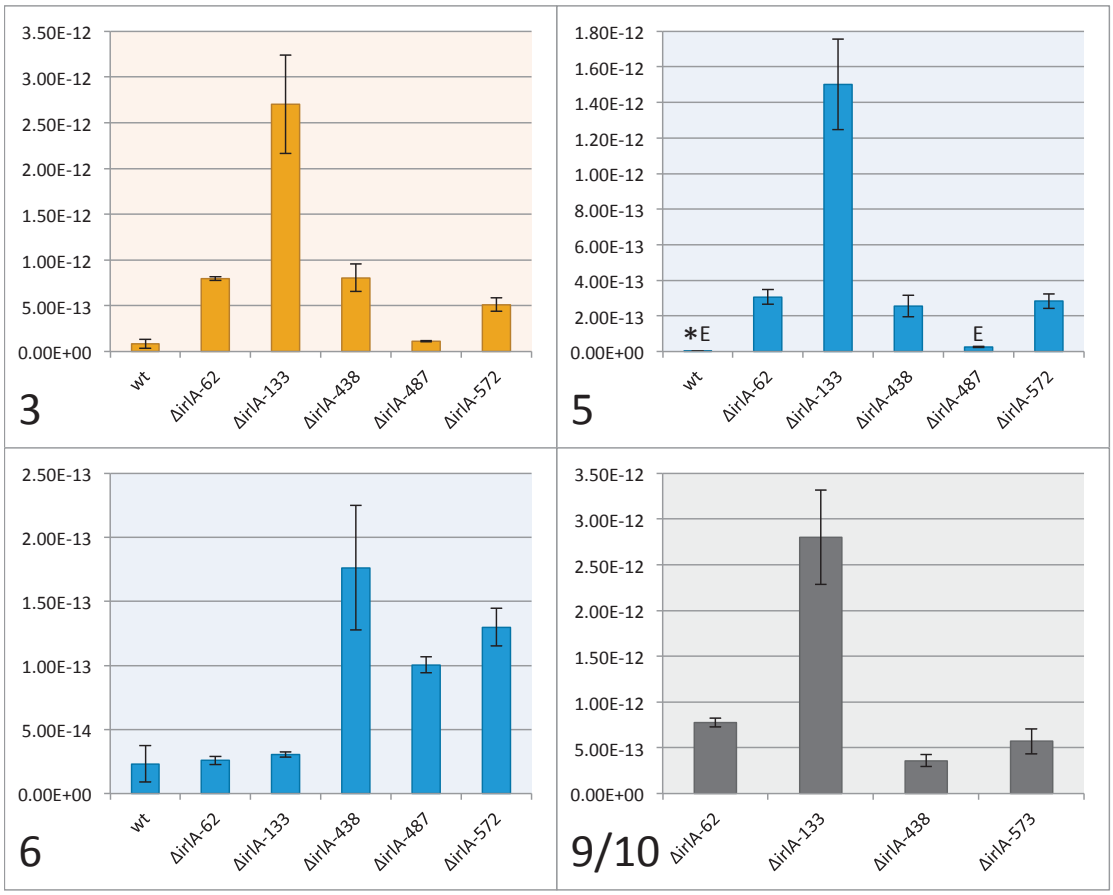


Figure 5.14 RT-qPCR analysis of *irlA* locus RNA expression in culture. Y-axis values represent the mean molar concentration of the target RNA across three biological replicates of each strain. The ID of the qPCR primer pair used to generate the data presented in each graph is shown in the bottom left corner (primers detailed in Table 5.2, see Figure 5.13 for visualisation of primer target sequences). Graph 9/10 was generated with data from primer pair 9 for the first two columns, with the second two columns generated using data from primer pair 10. Displayed values were normalised against expression of the reference gene *EF2*. Graphs of the same values normalised against expression of the reference gene *rps22* show similar results (data not shown). Error bars show the standard error of the mean for each strain/primer pair combination. The symbol E indicates that the calculated concentration of the associated sample mean is extrapolated beyond the lowest standard used for that primer pair (1×10^{-13} for primer pair 5 and 1×10^{-14} for all other primer pairs). The symbol * indicates that the target RNA was not detected in one or more of the technical triplicates used to generate the concentration of at least one of the biological replicates from the associated strain.

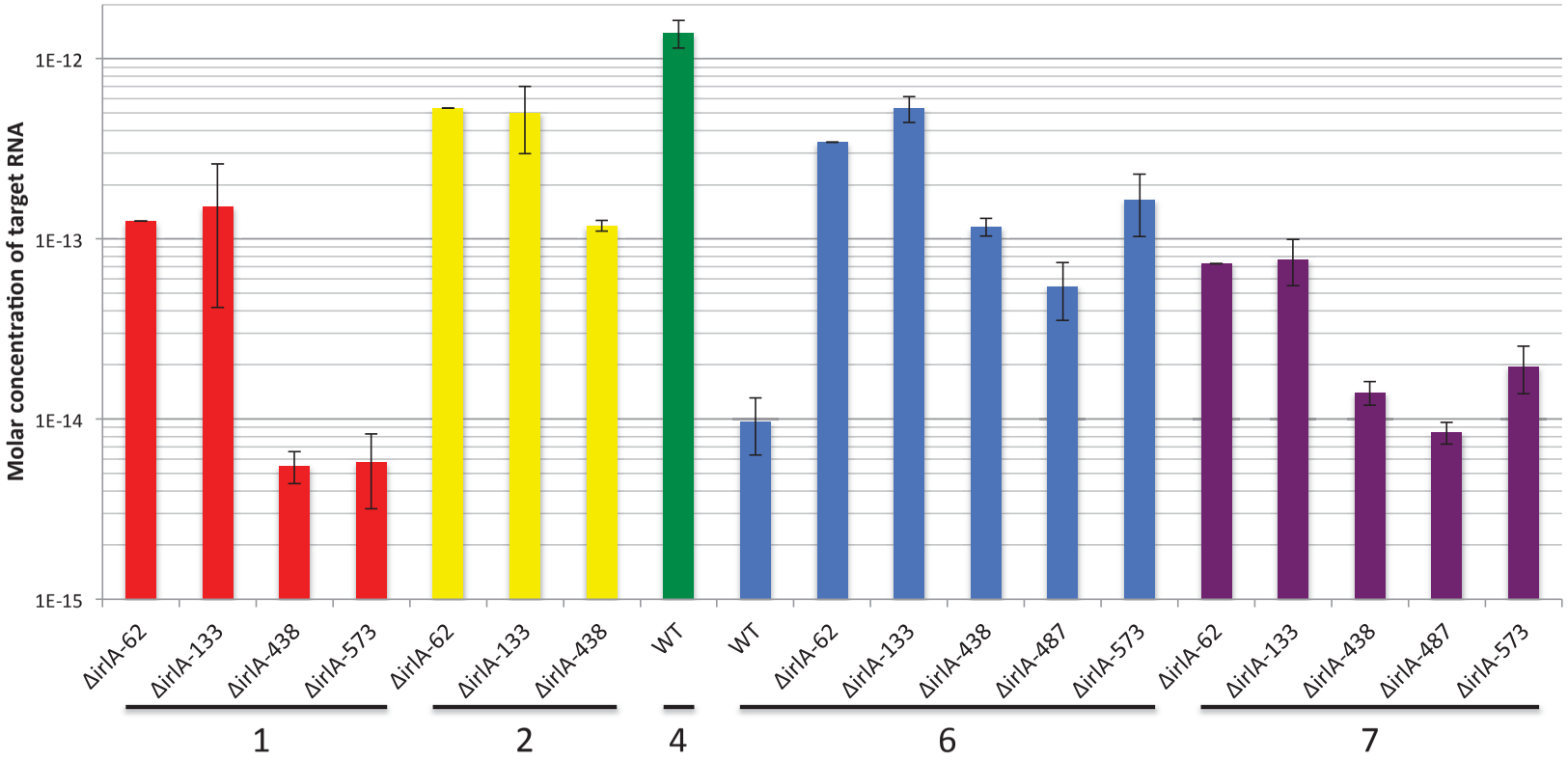


Figure 5.15 RT-qPCR analysis of *irlA* cluster gene expression for Δ*irlA* mutants *in planta*. The ID of the qPCR primer pair used to generate each column is annotated below the sample ID (primers detailed in Table 5.2; see Figure 5.13 for visualisation of primer target sequences). Values represent the mean molar concentration of the target RNA across three biological replicates of each strain, except Δ*irlA*-62, for which only one biological replicate was available. Strain/primer pair combinations that did not produce detectable results are not shown on this graph. These values are normalised against expression of the reference gene *EF2*. The lowest standard concentration used was 1×10^{-13} for primer pair 2 and 1×10^{-14} for all other primer pairs; values below these thresholds are extrapolated and therefore unreliable. Graphs of the same values normalised against expression of the reference gene *rps22* show similar results (data not shown). Error bars show the standard error of the mean for each strain/primer pair combination.

5.5 Discussion

Results presented in this chapter establish IrlA as being similar to PIP and isoflavone reductase-like (IRL) family proteins, with previous studies providing strong RNA-seq based support for a symbiosis-related role (Eaton et al., 2010; Eaton et al., 2015; Chujo et al., unpublished data). To date PIP and IRL-family proteins have been studied primarily in plants (Hua et al., 2013; Louie et al., 2007; Min et al., 2003; Wang et al., 2006). Interestingly, these PIP/IRL proteins all catalyse reactions involving isoflavanoid and lignan substrates that contain phenolic hydroxyl groups (Louie et al., 2007; Min et al., 2003; Wang et al., 2006), and this phenolic hydroxyl group has been shown to be required from substrate reduction by eugenol synthase (Koeduka et al., 2006). This suggests that the IrlA substrate may similarly contain a phenolic hydroxyl group. The large number of strong BLASTp matches observed on GenBank using an IrlA protein sequence query indicates that IRL proteins may also be highly prevalent within the filamentous fungi, but many hits are to "NmrA like" proteins. IFR/IRL proteins and NmrA proteins all derive from the short-chain dehydrogenase / reductase (SDR) family, and share the common Rossmann fold protein structural motif required for NAD/NADPH binding (Stammers et al., 2001; Wang et al., 2006). However, NmrA is catalytically inactive, instead functioning as a post-translational regulator of the nitrogen metabolism-associated transcription factor AreA (Stammers et al., 2001). The protein family database Pfam cautions that the NmrA-like family "only contains a few sequences as iteration results in significant matches to other Rossmann fold families" (Pfam, 2016. *NmrA-like family*. Retrieved from http://pfam.xfam.org/family/PF05368). Many of these "NmrA-like" proteins detected when searching with an IFL protein sequence query might therefore be better described as IFL proteins.

Deletion analysis of *irlA* was initially very promising, with a strong symbiosis-defective phenotype observed in four independent Δ*irlA* mutants. Previous deletions of *Epichloë* SM genes were predominately asymptomatic (Becker et al., 2015; Eaton et al., 2008; Tanaka et al., 2013; Tanaka et al., 2006), though deletion of the NRPS-encoding *sidN* gene did result in a symbiosis-defective phenotype (Johnson et al., 2013b). The failure to complement this Δ*irlA* symbiosis-defective phenotype by

ectopic reintroduction of *irlA* and the failure of pDB36-derived Δ*irlA* mutants to reproduce this phenotype cast doubt on the validity of this interpretation. The failure of a complementation experiment to rescue a mutant-associated phenotype is often interpreted as meaning that unanticipated mutations independent of the targeted gene deletion are the cause of the observed phenotype; however, in this case there were four independently generated pDB02-derived Δ*irlA* mutants exhibiting the same host-stunting phenotype. This essentially eliminates the possibility that random genetic background changes during or subsequent to the transformation are responsible for the observed phenotype.

An alternative hypothesis is that the protoplasts used to generate these mutants contained a systemic, pre-existing mutation causing this symbiotic phenotype. This is considered extremely unlikely as these protoplasts were sourced from a large general laboratory stock, and other users of these protoplasts encountered no similar complications. Additionally, the genes *hph* and *nptII* are the primary selectable markers used for *E. festucae* transformations in our lab, with all *E. festucae* mutants containing at least one of these markers. If protoplasts were accidentally used that had been generated from previous artificially generated gene deletions, the presence of the selectable marker *hph* (hygromycin resistance) within the protoplast background would have been quickly picked up by the transformation controls. Growth inhibition of Δ*irlA*-62 by geneticin was also tested and confirmed (data not shown). Together, these observations strongly suggest a causative relationship between the symbiosis-defective phenotype and insertion of the pDB02-derived *irlA* deletion construct at the target locus. While no pDB02-derived single-copy integration mutant was identified, such a mutant is expected to be genetically similar to Δ*irlA*-487, and would therefore likely also be asymptomatic *in planta*. The sequences of both the pDB02-derived and pDB36-derived deletion constructs were carefully verified, meaning that the only difference between these two constructs that could be responsible for this phenotypic difference is the additional 157 bp length of the right border in the pDB02-derived *irlA* deletion construct. Together, this implies that tandem integration of the pDB02-derived deletion construct is a requirement for this symbiosis-defective phenotype.

Right borders from pDB02-derived and pDB36-derived deletion constructs contain 742 bp and 585 bp fragments of the gene *mfsB* at their 3' ends, respectively. RT-qPCR analysis of the Δ*irlA* tandem integration mutants showed that transcription initiated by the promoter P*trpC*, which was intended to drive *hph* expression only, causes antisense transcription of the downstream *mfsB* gene fragment in tandem insertion mutants. One hypothesis that could explain the phenotypic difference between pDB02-derived and pDB36-derived tandem insertion Δ*irlA* mutants is that antisense *mfsB* RNA binds to complementary *mfsB* mRNA to form double-stranded RNA (dsRNA), resulting in post-transcriptional silencing of *mfsB* mediated by the RNA interference (RNAi) pathway. Previous studies have shown that RNAi-induced post-transcriptional gene silencing is effective in *Epichloë uncinatum* (Pan et al., 2014a; Spiering et al., 2005), indicating that the same is likely to be true for *E. festucae*. The target specificity of RNAi-based gene silencing is based on small interfering RNA (siRNA) molecules complementary to mRNA transcribed from the target gene (Wilson and Doudna, 2013). These siRNA are generated by the ribonuclease Dicer, which cleaves dsRNA into 21-23 bp fragments. Once converted to single stranded RNA these fragments are loaded onto the RNA-induced silencing protein complex (RISC), inducing cleavage of complementary mRNA molecules (Wilson and Doudna, 2013). It is well established that siRNA molecules complementary to different regions of the same gene can induce dramatically different levels of RNAi-mediated post-transcriptional gene silencing (Bohula et al., 2003; Holen et al., 2002; Luo and Chang, 2004). The pDB02-derived *irlA* deletion construct right border is 157 bp longer than the pDB36-derived *irlA* deletion construct. This would increase the length of antisense *mfsB* RNA generated by pDB02-derived tandem integration Δ*irlA* mutants from these extra right border *mfsB* fragments. This might enable greater induction of *mfsB*-targeted RNAi than mutants generated with the shorter pDB36-derived *irlA* deletion construct. This might be simply due to the additional length of antisense mRNA generated, or alternatively because this additional sequence results in production of an siRNA that is more efficient at inducing silencing. Due to the nature of cDNA synthesis using a combination of oligo(dT) and random hexamer primers it is not possible to differentiate between sense and antisense *mfsB* transcript. This means that it is not possible to directly confirm silencing of *mfsB* due

to antisense transcription from the results presented in this thesis. However, the observation that expression of RNA from the region of *mfsB* not located within the *irlA* deletion construct right border is raised in pDB36-derived Δ*irlA* mutants but not pDB02-derived Δ*irlA* mutants is suggestive of silencing (Figure 5.13). In a future experiment it may be possible to do a two-step reverse transcription reaction to obtain cDNA from the sense *mfsB* mRNA template only. The initial reverse transcription step would be done using primers specific to the sense *mfsB* mRNA only, and the subsequent PCR step would be done after RNase treatment and/or inactivation of the reverse transcriptase. This would allow specific amplification of sense *mfsB* mRNA. It is important to note that silencing of *mfsB* alone is unlikely to cause the observed phenotype, as in Section 6.4 it is shown that the Δ*mfsB* strain does not induce a symbiosis-defective phenotype *in planta*. It is therefore likely that this phenotype is caused by a combination of *irlA* deletion and *mfsB* silencing, although previous studies have suggested that in some cases RNAi-induced silencing of a gene can result in a more severe phenotype than gene deletion due to the activation of compensatory pathways (Rossi et al., 2015).

Another possibility for the phenotypic disparity between pDB02-derived and pDB36-derived tandem integration Δ*irlA* mutants is that off-target silencing of an unknown gene is induced by the extra 157 bp of antisense RNA transcribed in pDB02-derived mutants. However, the observed upregulation of the genes flanking *irlA* for pDB02-derived mutants *in planta* demonstrates that the pDB02-derived mutants do induce changes proximal to the Δ*irlA* locus. This suggests that off-target effects are not required to explain the phenotypic disparity between pDB02-derived and pDB36-derived mutants *in planta*. In Chapter 6 it is proposed that *irlA* is part of a five-gene SM cluster that drives a biosynthetic pathway similar to that of the characterised *lna* and *lnb* gene clusters from *Aspergillus flavus* (Forseth et al., 2013). The other genes in this pathway are *pdtA*, encoding a PLP-dependent transferase, *afrA*, encoding an adenylate-forming reductase, *mfsB*, encoding a major facilitator superfamily transporter, and *fxbA*, encoding an FAD-binding oxidase. A model for the proposed biosynthetic pathway of this cluster, based on both the *A. flavus lna*/*lnb* clusters (Forseth et al., 2013) and the results presented in this thesis, is shown in Figure 5.16

It is shown both in this and previous studies that, of the five genes in the *irlA* cluster, only *irlA* itself is expressed at significant levels during vegetative growth *in planta* (Eaton et al., 2010; Eaton et al., 2015; Chujo et al., unpublished data). However, in Chapter 6 it is shown that expression of the remaining cluster genes is specific to the highly proliferative pre-sexual stromata structures. A maintenance role is therefore proposed for *irlA*, ensuring that any product from low-level pathway throughput during vegetative growth *in planta* is converted to a non-proliferative end product. In the Δ*irlA* mutants IrlA is not available to fulfil this role, so any pathway throughput results in synthesis of the proliferation-promoting product. This could result in feedback upregulation of the remaining pathway genes, as was observed by Forseth et al. (2013). Once *mfsB* is expressed, MfsB would then transport this proliferation-promoting product out of the cell, with no consequences to the symbiosis due to the low levels produced. However, silencing of *mfsB* would result in product accumulation within the cell, promoting further feedback upregulation and proliferative growth normally only associated with stromata. While highly speculative, such a model would explain why the symbiosis-defective phenotype was only observed in pDB02-derived tandem integration mutants. Expression of the *irlA* cluster genes *pdtA*, *afrA* and *fxbA* was observed in all tandem integration Δ*irlA* mutants, but levels were significantly higher in pDB02-derived tandem integration Δ*irlA* mutants compared to the pDB36-derived tandem and single-copy integration mutants. This indicates a threshold at which the symbiosis-defective phenotype becomes apparent, which the pDB36-derived mutants do not reach due to MfsB preventing product accumulation (Figure 5.16).

Based on this model, a future experiment that could confirm involvement of the *irlA* deletion in the pDB02-derived Δ*irlA* phenotype *in planta* would be analysis of a double Δ*irlA*/Δ*mfsB* deletion mutant, which would be expected to exhibit the same *in planta* phenotype as the pDB02-derived tandem insertion mutants. It should be noted that this model does not explain why ectopic reintroduction of *irlA* failed to rescue the Δ*irlA*-62 symbiosis-defective phenotype, although involvement of RNAi could have been a complicating factor. It is possible that position-specific effects of these complementation mutants prevent sufficient transcription of *irlA* to restore wt-like growth. Screening additional complementation mutants, or complementing

with *irlA* under the control of a strong, constitutively active promoter may yet yield a complemented phenotype. Attempts were made to generate knock-in complementation strains from Δ*irlA*-62 and protoplasts, but were unsuccessful (data not shown). This was likely due to the presence of multiple copies of the deletion construct in this mutant, as knock-in of a complementation construct in the single-copy integration mutant Δ*irlA*-487 was successful, if ultimately unnecessary due to the lack of *in planta* phenotype to complement (data not shown).

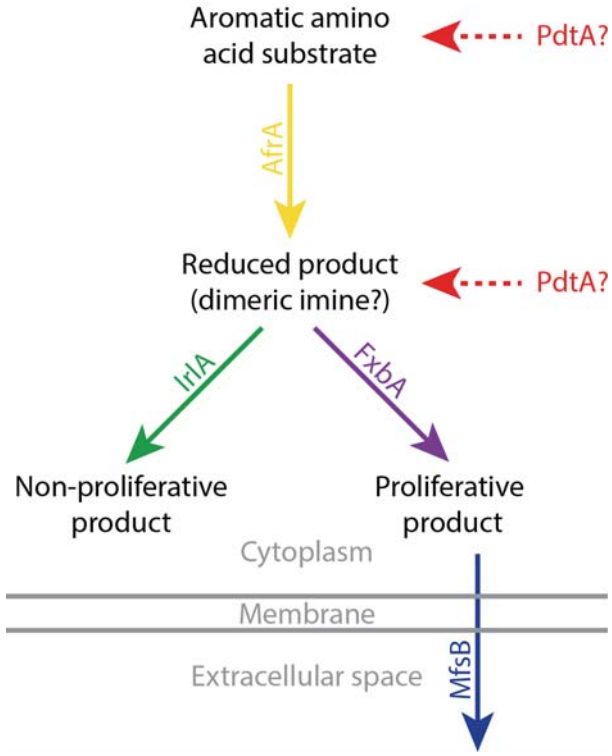


Figure 5.16 Model of biosynthetic pathway involving IrlA as described in Chapter 6.

The complications encountered in generating *irlA* gene deletion mutants suggest a number of improvements that could be made in future to avoid similar occurrences. The deletion construct in this experiment was designed without a heterologous terminator following *hph*. Excluding the terminator results in a smaller *hph* expression cassette, which can help the number of false negatives when PCR amplifying across the expression cassette from crude gDNA template. PCR amplification across the *hph* expression cassette is commonly used for differentiating between *E. festucae* transformants containing targeted and ectopic

integrations of the deletion construct. The *irlA* deletion constructs were designed so the native T*irlA* would terminate P*trpC*-driven transcription of *hph*; however, the results presented here show that this was not sufficient. This may be due to the omission of the first 15 bp of T*irlA* due to primer design considerations, but it would nevertheless be prudent for future deletion constructs to include a strong terminator following the selectable marker gene.

Overall 600 strains transformed with the pDB02- or pDB36-derived *irlA* deletion constructs were screened, from which ten Δ*irlA* strains were identified, representing a target insertion frequency of 1.67%. Such low efficiencies for targeted insertions are not unusual for *Epichloë festucae*, but efficiencies in excess of 50% have been reported for some loci (K. Green, personal communication). High locus-dependent variability in homologous recombination-based integrations has been observed in other filamentous fungi (Bird and Bradshaw, 1997), but the mechanism behind this phenomenon is unclear. Expression levels of the target gene in culture do not appear to influence the integration frequency (Bird and Bradshaw, 1997). However, other factors such as the chromatin state at the target locus or presence of nearby sequences that promote or inhibit homologous recombination may influence targeted integration frequency. Significant problems were encountered obtaining single-copy integration Δ*irlA* mutants, with 600 transformants screened to identify a single example. It is possible that deletion construct concatamers may be more efficient than single units at integrating into regions for which homologous recombination is inefficient. This could be because the presence of multiple homologous flanking sequences in a concatamer increases the probability of both flanks interacting and recombining with the target sequences, as required for a double crossover integration event (Krappmann et al., 2006).

Changing the restriction enzymes used to excise the deletion constructs so that incompatible sticky ends were generated caused no obvious difference to tandem integration frequency. It is possible that using blunt-ended DNA fragments generated by restriction enzyme, or by PCR using a DNA polymerase with a proofreading exonuclease, might inhibit deletion construct concatemerisation and thus reduce tandem integration frequency. However, the use of PCR amplification products should be carefully managed to avoid the potential for PCR-based SNPs to

cause unanticipated deleterious mutations. A more effective strategy might be to utilise an advanced method for targeted gene deletion, such as split-marker transformation of protoplasts (Fairhead et al., 1996; Gravelat et al., 2012), use of a two-marker strategy that selects against integrations that occur via the non-homologous end-joining mechanism (Gardiner and Howlett, 2004; Liu et al., 2001), or using the highly effective CRISPR gene editing technology to induce double stranded breaks in the target gene, dramatically increasing the homologous recombination frequency at that location (Liu et al., 2015; Nodvig et al., 2015).

6 *IrlA* is part of a five gene secondary metabolism gene cluster linked to stroma formation

6.1 Results: synteny analysis defines a five-gene cluster centred on *irlA*

Synteny analysis of the *irlA* locus across the Clavicipitaceae identified a cluster of five consecutive genes centred on *irlA* in which the gene sequence, order and orientation were conserved between all sexual *Epichloë* spp., but were not present at this locus in any of the other Clavicipitaceaen genomes analysed (Figure 6.1). This cluster was also absent in the genomes of *Epichloë gansuensis* and *Epichloë inebrians*, which are asexual, non-hybrid species. The identity and order of those genes located flanking the *Epichloë irlA* cluster were highly conserved across the Clavicipitaceae (Figure 6.1; Figure 6.2). The genes within this predicted cluster showed some overlap with the symbiotically regulated nitrogen metabolism cluster predicted by Eaton et al. (2010).

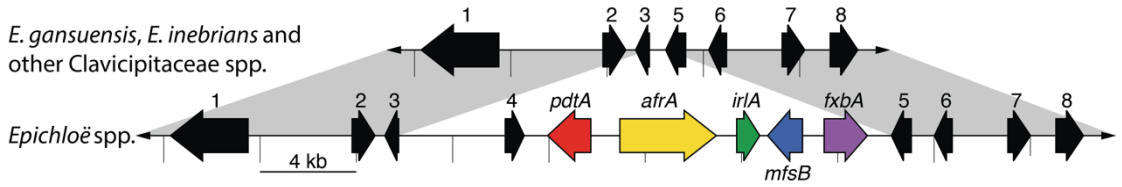


Figure 6.1 Synteny analysis of the *irlA* locus across the Clavicipitaceae. All available Clavicipitacean genomes were analysed by tBLASTn using the encoded protein sequence predicted for each of the *E. festucae* Fl1 *irlA* cluster genes, as well as the eight closest genes flanking this cluster. The *irlA* cluster gene names are defined in Section 5.5. This figure shows a generalised representation of the *irlA* locus for *Epichloë* spp. and other Clavicipitaceae; the structure of this locus within individual genomes may vary slightly from the gene map presented here.

6.2 Defining putative functions for *irlA* cluster gene products

Multiple sequence alignments between all available *Epichloë* sequences of each *irlA* cluster gene were used to confirm or refine the transcription start and stop sites of the putative *E. festucae* Fl1 gene models (Schardl et al., 2013a). RNA-seq data from wt infected *L. perenne* plants (Eaton et al., 2010) was used to define the

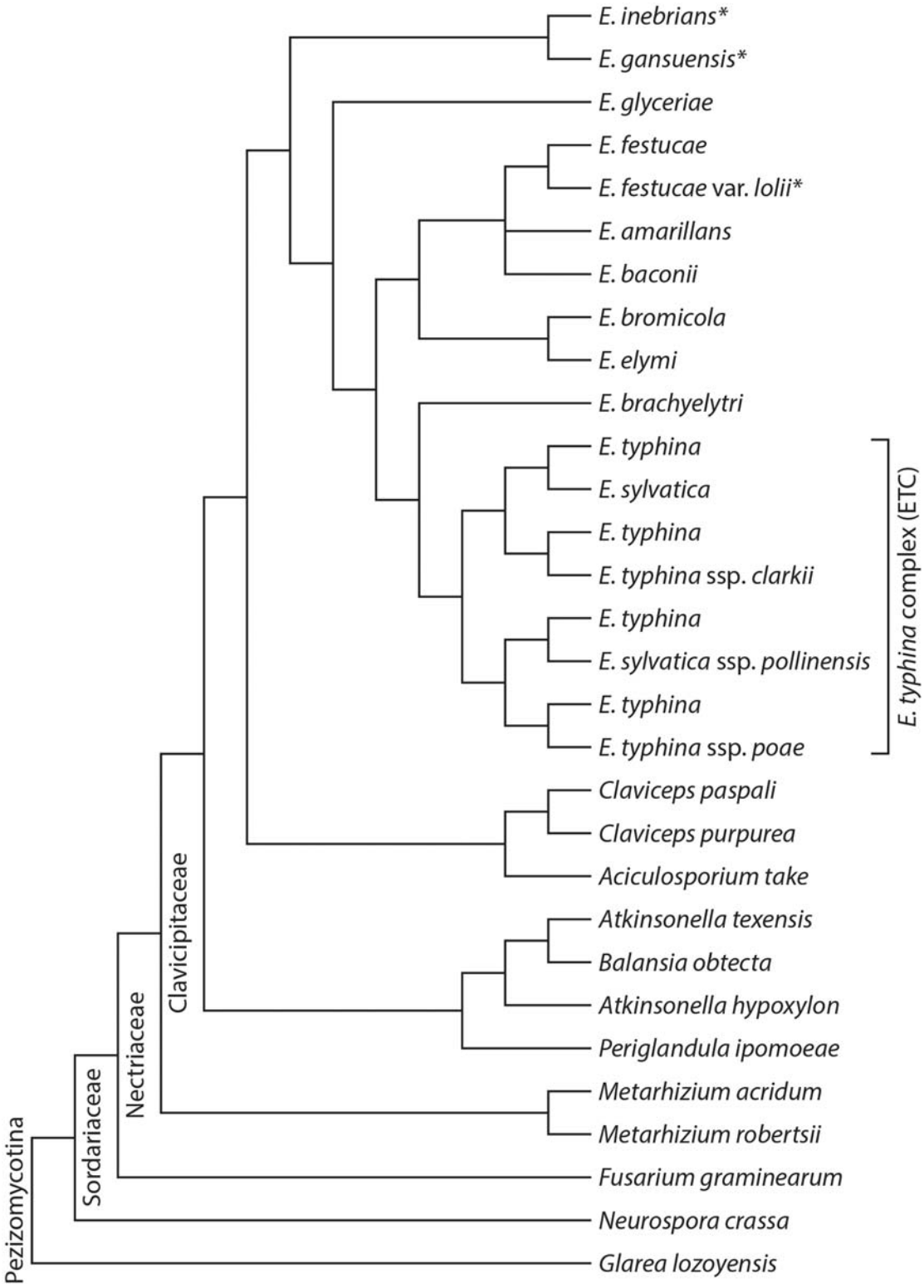


Figure 6.2 Cladogram approximating the taxonomy of family Clavicipitaceae. Cladogram is a simplification of the results presented by Schardl et al. (2014). Asexual *Epichloë* spp. are indicted with an asterisk (*).

intron locations within these genes. The predicted protein sequences encoded by each gene were then used as queries for the InterPro online search tool (Jones et al., 2014) to assign putative domains and functions for each protein. Two introns were identified in the first gene of this cluster, which was predicted to encode a 547 aa protein containing a pyridoxal phosphate (PLP)-dependent transferase domain (Figure 6.3). This protein was named PdtA, for **P**LP-**D**ependent **T**ransferase, encoded by the gene *pdtA*.

No introns were identified in the second gene, which was predicted to encode a 1334 aa protein containing AMPylation, PCP and reductase domains (Figure 6.3). This domain architecture is consistent with an adenylate-forming reductase (AFR) family protein (Kalb et al., 2014), which are closely related to non-ribosomal peptide synthetases (NRPS). This protein was therefore named AfrA, encoded by the gene *afrA*. NRPSpredictor2 determined that the 10AA substrate-binding code of the AfrA A-domain was DIHDIAAICK (Roettig et al., 2011). While substrate prediction for fungal A-domains is significantly less accurate than bacterial A-domains due to a relative dearth of fungal domains with characterised substrates, analysis of the AfrA A-domain sequence using NRPSpredictor2 did predict that the substrate was likely to be an amino acid with a large hydrophobic sidechain (Roettig et al., 2011). The 10AA code of AfrA also shows striking similarity to that of the phenyalanine-binding bacterial A-domain PheA (DAWTIAAICK; Stachelhaus et al., 1999), as well as many fungal phe-binding A-domains (Kalb et al., 2013; Roettig et al., 2011). This suggests that the substrate of AfrA may be an aromatic amino acid such as phenylalanine or tyrosine. AfrA was also found to contain a 250 aa N-terminal extension to the A-domain with predicted structural similarity to the second subunit of an NRPS C-domain (Kelley et al., 2015). Based on the results of a previously published study, this extension is likely to be a requirement for A-domain function (Kalb et al., 2015).

The fourth gene in this cluster, located immediately downstream of *irlA*, was predicted to encode a 463 aa major facilitator superfamily (MFS) transporter protein. This protein was therefore named MfsB, encoded by *mfsB*, and contained 12 predicted transmembrane domains, as is typical for MFS transporters (Figure 6.3). A large polypeptide loop was identified separating the 6th and 7th transmembrane

domains in MfsB, with this loop predicted to be located on the cytoplasmic side of the membrane.

Two introns were identified in the final gene, which encoded a predicted 551 aa protein containing a FAD-binding domain and a C-terminal berberine bridge enzyme-like domain (Figure 6.3). This protein was predicted to be an oxidoreductase, and was therefore named FxbA for **F**AD-binding o**X**idoreductase with **B**erberine bridge enzyme-like domain. A single N-terminal transmembrane domain was also predicted for FxbA, indicating it may be a membrane-tethered protein, with the protein sequence downstream of this transmembrane domain predicted to be cytoplasmic (Figure 6.3).

6.3 Identification of potential *irlA* cluster gene orthologs in other Clavicipitacean genera

All available genomes from Clavicipitaceae genera besides *Epichloë* were interrogated by BLASTn and tBLASTn using the CDS or predicted encoded protein sequence from *E. festucae* Fl1 *pdtA*, *afrA*, *irlA*, *mfsB* and *fxbA* (Table 6.1; Table 6.2). These sequences were also used to back-query the *E. festucae* Fl1 genome. It was predicted that the E-value of a BLAST match between an *Epichloë festucae* Fl1 gene query and an orthologous sequence from another Clavicipitacean spp. should be higher than the E-value of the second best match in the *E. festucae* Fl1 genome, assuming no gene duplication events (Table 6.2; Table 6.1). Therefore, BLAST matches to other Clavicipitacean genomes using *irlA* cluster queries were disregarded as potential orthologs if they produced E-values that were not significantly stronger than the second best *E. festucae* Fl1 match. This method was tested by using *irlA* cluster gene queries against the genomes of *E. gansuensis* and *E. inebrians*, which synteny analysis had shown did not contain the *irlA* cluster at the same locus as in other *Epichloë* spp. (Figure 6.1). The E-values from this analysis were similar to the E-values from the second-best *E. festucae* matches, indicating that the *irlA* cluster genes were entirely absent from the *E. gansuensis* and *E. inebrians* genomes (Table 6.2; Table 6.1). Because they appeared to lack the *irlA* cluster, the E-values from the best matches against the genomes of *E. gansuensis*

and *E. inebrians* were also used to filter results from queries against other Clavicipitacean genomes. "Strong" matches were therefore defined where the BLASTn E-value for a match was substantially better (lower) than the second-best match against the genome of *E. festucae* Fl1 and best matches against the genomes of both *E. gansuensis* and *E. inebrians*. Sequences identified as strong matches were then used as queries for reciprocal BLASTn against the *E. festucae* Fl1 genome to double-check that these sequences were true homologs of the *irlA* cluster genes.

6.3.1 The five-gene *irlA* cluster is present in *Aciculosporium take*

The only Clavicipitacean genome outside of *Epichloë* for which BLASTn searches identified orthologs of all five *E. festucae irlA* cluster genes was the bamboo endophyte *A. take*, in which all five top hits co-localised to a single contig (Table 6.1). The sequence encoded by that contig was not proximal to the locus in which the *irlA* cluster is located in *Epichloë* spp. However, the *pdtA*, *afrA*, *irlA*, *mfsB* and *fxbA* homologues in *A. take* were clustered together in the same order and orientation as in the *Epichloë irlA* cluster (Figure 6.4). Reciprocal BLASTn analysis using the predicted CDS from these each putative *A. take irlA* cluster gene as queries against the *E. festucae* Fl1 genome confirmed that the equivalent *E. festucae irlA* cluster genes were the strongest match in every case [strongest (lowest) E-values were 4×10^{-24} for *pdtA*, 5×10^{-38} for *afrA*, 2×10^{-24} for *irlA*, 1×10^{-32} for *mfsB* and 9×10^{-9} for *fxbA*; second-best E-values were all weaker (higher) than 1×10^{-2}]. Analysis of these orthologous sequences from *A. take* did not identify any premature stop codons, frameshift events, large indels or other genetically disruptive events. This suggested that these homologous *A. take* sequences encoded functional orthologs of *Epichloë pdtA*, *afrA*, *irlA*, *mfsB* and *fxbA*.

6.3.2 Degenerate *irlA* clusters are present in *Metarhizium* spp.

Strong BLASTn matches were observed for the genomes of *Metarhizium acridum* and *Metarhizium robertsii*, with *M. acridum* exhibiting co-localised matches to *pdtA* and *fxbA*, and *M. robertsii* exhibiting a single strong match to *fxbA* (Table 6.1). Reciprocal BLAST analysis of these *Metarhizium* sequences against the *E. festucae* Fl1 genome found that the *E. festucae irlA* cluster genes were the closest matches in every case, with corresponding E-values of 5×10^{-5}, 6×10^{-11} and 7×10^{-7}

against the *M. acridum pdtA*, *fxbA* and *M. robertsii fxbA* sequence queries, respectively (all other E-values were >1×10^{-2}). Closer analysis of the sequences surrounding these matches revealed that *M. robertsii* contained a degenerated cluster that was more complete than the BLASTn analysis has indicated, with gene fragments homologous to *mfsB* (668 bp with 60% sequence identity to *E. festucae mfsB*) and *fxbA* (400 bp with 60% sequence identity to *E. festucae fxbA*; Figure 6.4). The degenerate cluster contained within *M. acridum* was also more complete than BLASTn analysis indicated, containing gene fragments homologous to *pdtA*, *afrA*, *irlA* and an apparently intact *fxbA* ortholog (fragment length/sequence identitiy to homologous Fl1 genes of 401 bp/67%, 1469 bp/63%, 376 bp/59% and 1812/61%, respectively; Figure 6.4). The order and orientation of the *M. acridum* gene fragments was also well conserved with the *irlA* cluster from *A. take*/*Epichloë* spp., with the only differences being the absence of *mfsB* and an inversion of *irlA* (Figure 6.4). It is unclear why the BLASTn approach failed to identify these additional homologous gene fragments in the *Metarhizium* genomes, and this suggests that undetected *irlA* cluster pseudogenes might also be present in the genomes of other Clavicipitaceaen species. Analysis of the sequences flanking the *irlA* clusters of both *M. acridum* and *M. robertsii* identified multiple genes that were homologous between these two species, which indicated that they share the same genomic locus (Figure 6.4).

6.4 A homologous gene cluster is present in the Leotiomycete *Glarea lozoyensis*

Searches using BLASTn and tBLASTn failed to identify obvious homologs of any *irlA* cluster genes outside of the Clavicipitaceae, with the sole exception of the Leotiomycete *Glarea lozoyensis*. A full five-gene cluster homologous to the *irlA* cluster of *E. festucae* was identified in *G. lozoyensis* (Figure 6.5A). Reciprocal BLASTn analysis using the predicted CDS from these each putative *G. lozoyensis irlA* cluster gene as queries against the *E. festucae* Fl1 genome confirmed that the equivalent *E. festucae irlA* cluster genes were the strongest match in every case [strongest (lowest) E-values were 5×10^{-30} for *pdtA*, 1×10^{-44} for *afrA*, 5×10^{-19} for *irlA*, 1×10^{-17} for

mfsB and 3×10^{-25} for *fxbA*; second-best E-values were all weaker (higher) than 1×10^{-2}]. The order and orientation of the genes in this *G. lozoyensis* cluster showed only a single inversion event observed relative to the *irlA* clusters of *A. take*/*Epichloë* spp. (Figure 6.4, Figure 6.5A). The genomic locus of this cluster was different to the *irlA* cluster loci previously described for *A. take*, *Epichloë* spp. and *Metarhizium* spp. (Figure 6.4, Figure 6.5A). The sequences of the *G. lozoyensis irlA* cluster gene homologs were unusually well conserved with those of both *E. festucae* and *A. take*, with the genes from each species showing very similar levels of sequence identity despite *G. lozoyensis* being relatively distantly related to the Clavicipitaceae (Figure 6.5B). *G. lozoyensis* contained predicted frameshift mutations within its *irlA* and *mfsB* homologs, and were therefore labelled as putative pseudogenes.

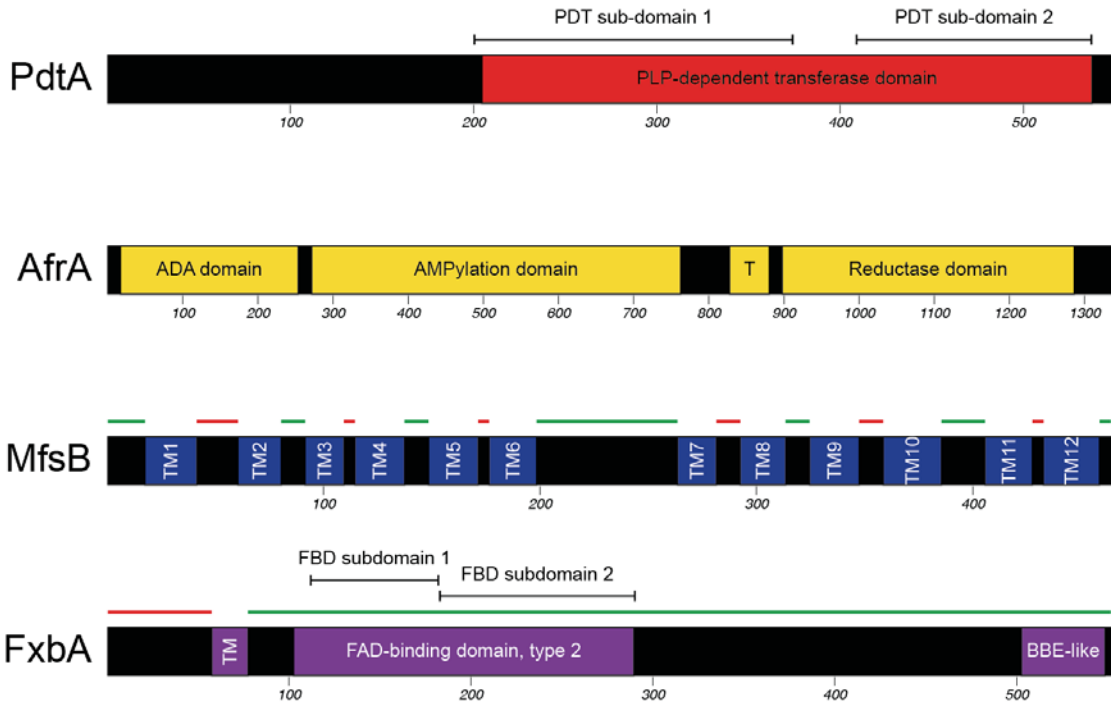


Figure 6.3 Predicted domains of PdtA, AfrA, MfsB and FxbA. Protein sequence maps of PdtA, AfrA, MfsB and FxbA are annotated with the protein domains predicted by InterPro scan analyses (Jones et al., 2014). The green and red bars above MfsB and FxbA indicate regions predicted to be cytoplasmic or periplasmic, respectively, assuming localisation to the cell membrane. Abbreviations are as follows: PLP = pyridoxal phosphate, PDT = PLP-dependent transferase, PCP = peptide carrier protein domain, TM = transmembrane domain, FAD = flavin adenine dinucleotide, FBD = FAD-binding domain, BBE = berberine bridge enzyme-like domain.

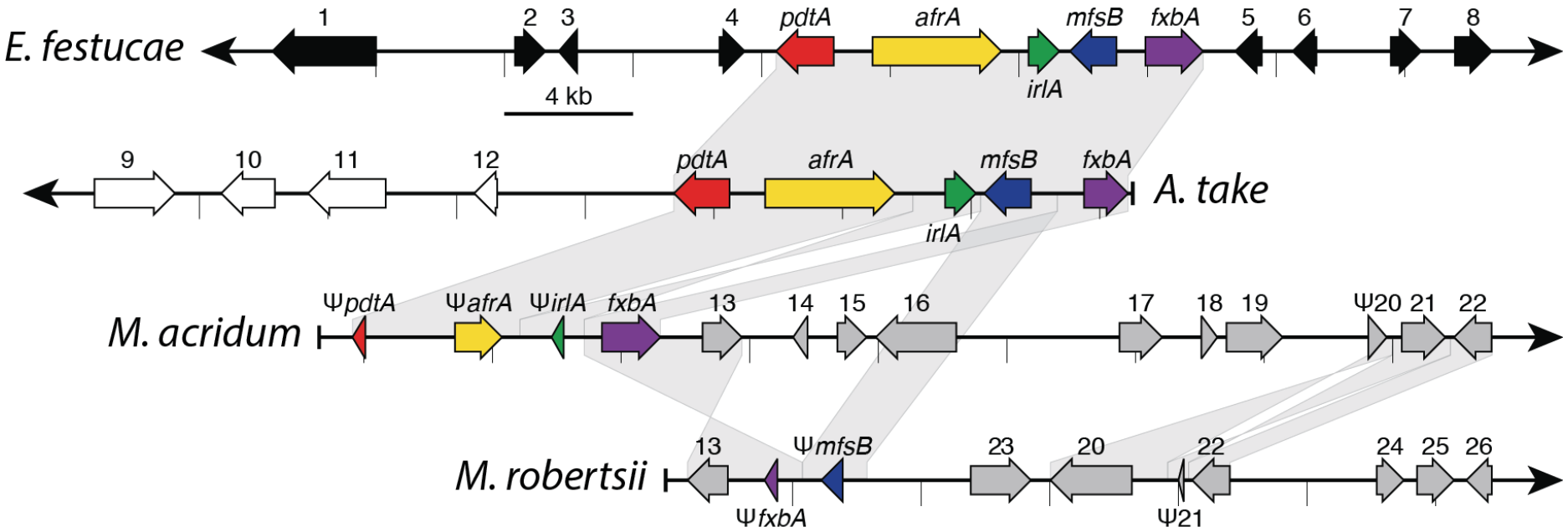


Figure 6.4 ***A. take* and *Metarhizium* spp. contain *irlA* cluster homologs.** Gene maps of the *irlA* cluster loci from *Epichloë festucae*, *Aciculosporium take*, *Metarhizium acridum* and *Metarhizium robertsii.* Grey drop-shadows indicate syntenic regions between species, while genes from different species annotated with the same name or number are homologous. The prefix Ψ identifies pseudogenes.

Table 6.1 Top E-values from BLASTn searches using *E. festucae* Fl1 *irlA* cluster gene queries

Subject genome		Sequence similarity to Fl1 query (strongest E-value)					Contig # associated with strongest E-value				
Species	**Strain**	***pdtA***	***afrA***	***irlA***	***mfsB***	***fxbA***	***pdtA***	***afrA***	***irlA***	***mfsB***	***fxbA***
Aciculosporium take	A7993	7×10^{-24}	8×10^{-38}	4×10^{-24}	2×10^{-32}	2×10^{-8}	00194	00194	00194	00194	00194
Atkinsonella hypoxylon	B4728	1.9×10^{-1}	8×10^{-6}	1.1×10^{-1}	1×10^{-2}	3	00174	00044	00064	00012	00230
Atkinsonella texensis	B6155	2×10^{-4}	4×10^{-4}	3.4×10^{-1}	8×10^{-3}	1×10^{-2}	00072	00054	00073	00028	00045
Atkinsonella texensis	B6156	2×10^{-4}	4×10^{-4}	3.4×10^{-1}	8×10^{-3}	1×10^{-2}	00016	00089	00014	00048	00040
Balansia obtecta	B249	1×10^{-8}	2×10^{-3}	9.3×10^{-2}	9×10^{-3}	1.6×10^{-1}	00186	00160	00060	00165	00327
Claviceps fusiformis	C3307	1.8×10^{-2}	3×10^{-3}	3×10^{-6}	5.9×10^{-2}	2.8×10^{-1}	00095	00495	00030	00010	00172
Claviceps paspali	C7990	7×10^{-7}	1×10^{-4}	2.2×10^{-2}	5.1×10^{-1}	1×10^{-2}	00755	00254	00922	00217	00117
Claviceps purpurea	C20x1	2.7	7×10^{-3}	2.5×10^{-2}	5.7×10^{-1}	1.7×10^{-1}	00110	00019	00006	00102	00018
Metarhizium acridum	CQMa102	2×10^{-4}	8×10^{-3}	4.6×10^{-1}	6.7×10^{-1}	6×10^{-11}	01313	00644	00458	00083	01313
Metarhizium robertsii	ARSEF23	5.2×10^{-2}	9×10^{-9}	8×10^{-3}	6.9×10^{-1}	4×10^{-6}	00443	00873	00522	00572	00781
Periglandula ipomoeae	p4806	5×10^{-8}	8×10^{-6}	4.3×10^{-1}	4×10^{-2}	7.5×10^{-1}	00403	00082	00898	00408	00111
Epichloë inebrians	E818	6×10^{-7}	1×10^{-4}	3.5×10^{-1}	5.1×10^{-1}	1.5×10^{-1}	00006	00286	00458	00402	00082
Epichloë gansuensis	E7080	2×10^{-4}	1×10^{-4}	4.8×10^{-1}	7×10^{-1}	2.1×10^{-1}	00004	00054	00052	00149	00525
Epichloë festucae	Fl1	0	0	0	0	0	00025	00025	00025	00025	00025
Epichloë festucae	Fl1 (2nd)[1]	2×10^{-4}	5×10^{-4}	1.1×10^{-1}	1×10^{-2}	4.7×10^{-2}	00001	00058	00009	00026	00036

Strong E-values are highlighted in dark green. Red highlights show where the top E-values for more than one gene query are localised to the same contig of a specific genome.

[1] Second-strongest match against the genome of *E. festucae* Fl1, used as a minimum-similarity reference when identifying putative homologs.

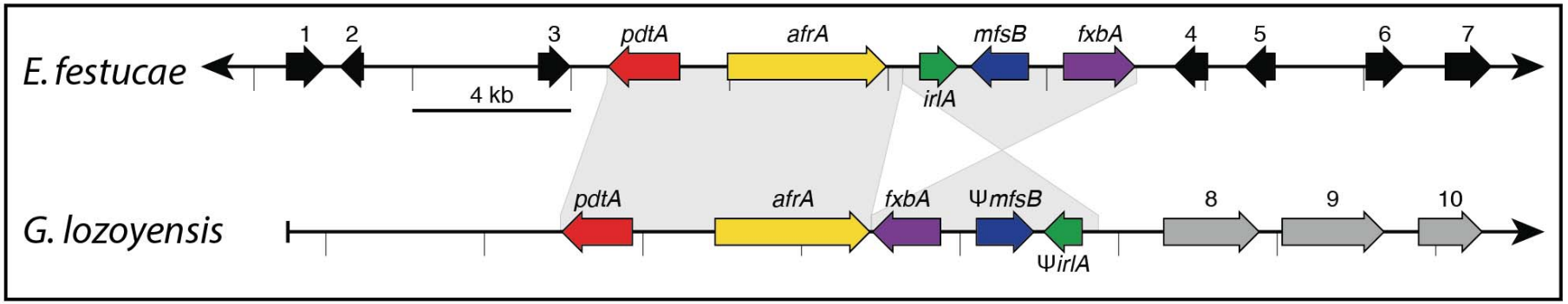
Table 6.2 Top E-values from tBLASTn searches using *E. festucae* Fl1 *irlA* cluster protein seqeunce queries

Subject genome		Sequence similarity to Fl1 query (strongest E-value)					Contig number associated with strongest E-value				
Species	**Strain**	**PdtA**	**AfrA**	**IrlA**	**MfsB**	**FxbA**	**PdtA**	**AfrA**	**IrlA**	**MfsB**	**FxbA**
Aciculosporium take	A7993	0	0	1×10^{-118}	0	1×10^{-165}	00194	00194	00194	00194	00194
Atkinsonella hypoxylon	B4728	2×10^{-65}	1×10^{-148}	6×10^{-10}	2×10^{-37}	2×10^{-36}	00008	00090	00096	00060	00030
Atkinsonella texensis	B6155	5×10^{-69}	1×10^{-139}	2×10^{-19}	1×10^{-33}	2×10^{-37}	00831	00017	00714	00779	00873
Atkinsonella texensis	B6156	5×10^{-69}	1×10^{-139}	2×10^{-19}	2×10^{-36}	2×10^{-37}	00324	00062	00763	00958	00871
Balansia obtecta	B249	5×10^{-74}	1×10^{-130}	4×10^{-21}	5×10^{-39}	8×10^{-43}	00186	00115	00117	00321	00003
Claviceps fusiformis	C3307	1×10^{-109}	0	3×10^{-75}	1×10^{-120}	7×10^{-32}	00095	00141	00030	00141	00483
Claviceps paspali	C7990	1×10^{-111}	1×10^{-115}	1×10^{-70}	1×10^{-34}	4×10^{-27}	00755	00377	00922	01249	00322
Claviceps purpurea	C20x1	1×10^{-110}	1×10^{-144}	7×10^{-76}	3×10^{-35}	6×10^{-29}	00093	00858	00902	00797	00134
Metarhizium acridum	CQMa102	1×10^{-107}	1×10^{-143}	2×10^{-86}	5×10^{-37}	1×10^{-152}	00049	00702	01021	00656	01313
Metarhizium robertsii	ARSEF23	1×10^{-102}	1×10^{-170}	9×10^{-87}	2×10^{-63}	4×10^{-61}	00443	00873	00522	00781	00568
Periglandula ipomoeae	p4806	9×10^{-77}	1×10^{-158}	4×10^{-75}	2×10^{-30}	3×10^{-39}	00403	00294	00184	00439	00002
Epichloë inebrians	E818	1×10^{-76}	1×10^{-145}	5×10^{-7}	9×10^{-36}	9×10^{-33}					
Epichloë gansuensis	E7080	1×10^{-108}	1×10^{-147}	6×10^{-6}	8×10^{-35}	6×10^{-31}					
Epichloë festucae	Fl1	1×10^{-149}	0	1×10^{-159}	1×10^{-146}	0	00025	00025	00025	00025	00025
Epichloë festucae	Fl1 (2nd)[1]	1×10^{-111}	1×10^{-169}	1.8×10^{-1}	7×10^{-55}	4×10^{-32}	00001	00013	00016	00026	00056

Strong E-values are highlighted in dark green and intermediate strength E-values are highlighted in light green. Red highlights show where the top E-values for more than one gene query are localised to the same contig of a specific genome.

[1] Second-strongest match against the genome of *E. festucae* Fl1, used as a minimum-similarity reference when identifying putative homologs.

A



B

	% identity of gene DNA sequence				
	pdtA	*afrA*	*irlA*	*mfsB*	*fxbA*
E. festcuae vs. *A. take*	71	69	72	72	65
E. festucae vs. *G. lozoyensis*	73	70	73	70	72
A. take vs. *G. lozoyensis*	68	68	70	71	69

Figure 6.5 Comparison of the homologous *irlA* gene clusters from *E. festucae* and *G. lozoyensis*. (A) Gene maps of the *irlA* cluster loci in *Epichloë festucae* and *Glarea lozoyensis.* Grey drop-shadows indicate syntenic regions between species. The prefix Ψ indicates pseudogene locations. (B) Table comparing the DNA sequence % identity from alignments of the *irlA* cluster genes from *E. festucae*, *A. take* and *G. lozoyensis*.

6.5 The *irlA* gene cluster is conserved in sexual *Epichloë* spp. only

The *irlA* cluster was manually annotated onto 24 different *Epichloë* genome sequences (Johnson et al., unpublished data; Schardl et al., 2013a; Schardl et al., 2013b), including six asexual non-hybrid strains and 15 strains from 10 different sexual species. The integrity of all five *irlA* cluster genes was fully conserved in all sexual strains, whereas four of the six asexual non-hybrid strains analysed contained small frameshift-causing insertions/deletions in one of the five genes or, in the case of *E. gansuensis* and *E. inebrians*, were missing the entire *irlA* gene cluster (Figure 6.6).

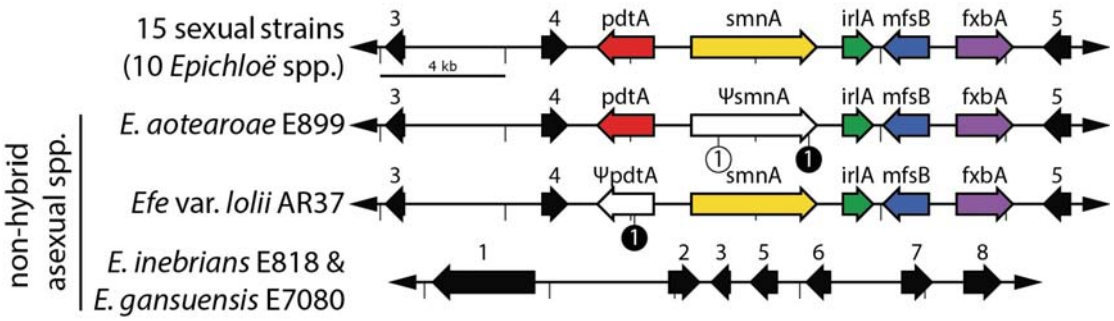


Figure 6.6 Conservation of *irlA* cluster genes is not maintained in asexual *Epichloë* spp. The top gene map is a generalised representation of the *irlA* locus in 15 strains covering 10 sexual *Epichloë* spp. Gene maps of the *irlA* locus from available genomes of asexual non-hybrids containing pseudogenes or gene deletions within the *irlA* cluster are shown below. Not shown here are the asexual non-hybrid strains *E. festucae* (*Efe*) var. *lolii* AR1 and PRG-3, which did not contain pseudogenes or deletions within the *irlA* cluster. The symbol Ψ identifies pseudogenes, with attached white or black-filled circles indicating the location of insertions or deletions, respectively, causing these pseudogenes; the numbers within these circles show the number of bp that have been inserted/deleted. Genes 1, 2, 6, 7 and 8 are present but not included in the first three gene maps due to spatial constraints. The sexual strains represented by the first gene map include: *E. amarillans* E57 and E4668; *E. baconii* E1031; *E. brachyelytri* E4804; *E. bromicola* AL0426/2, AL0434 and E502; *E. elymi* E56; *E. festucae* E2368 and Fl1; *E. glyceriae* E277; *E. mollis* E3601; *E. sylvatica* E7368; *E. typhina* E8 and *E. typhina* ssp. *poae* E5819. These gene maps were somewhat simplified for viewing clarity; intergenic region lengths varying between strains more than is indicated here, and gene 4 is not present at this locus in *E. brachyelytri*, *E. gansuensis* and *E. typhina* complex isolates.

6.6 The *irlA* gene cluster is upregulated in *Epichloë* stromata

Results from an RNA-seq comparison between choked and asymptomatic inflorescences of *E. festucae* E2368-infected *Festuca pratense* plants (tall fescue) were provided by Prof. Chris Schardl (University of Kentucky). These results indicated that expression of *irlA* cluster genes was dramatically upregulated in stromata (Appendix 9.5.1). Further experiments were therefore designed to determine if this expression pattern was consistent across multiple endophyte-host associations involving sexual *Epichloë* spp. Samples were taken from *Dactylis glomerata* (orchardgrass) plants near Corvallis, OR, USA displaying *E. typhina* stromata and *Elymus* sp. plants near Ardmore, OK, USA displaying *E. elymi* stromata. Three *D. glomerata* plants were sampled, with three tillers exhibiting fertilised stromata and three tillers exhibiting unfertilised stromata taken from each plant (Figure 1.2). These tillers were divided into basal, sub-stromata and stromata sections, which were then immediately frozen in liquid nitrogen. Only two *Elymus* sp. plants exhibiting stromata could be located. One tiller exhibiting an unfertilised stroma and one asymptomatic flowering tiller was taken from each *Elymus* sp. plant. These *Elymus* sp. tillers were divided into basal/stroma sections for the stromata-exhibiting tillers and basal/inflorescence sections for the asymptomatic tillers, then immediately frozen in liquid nitrogen. Asymptomatic inflorescences could not be sampled from *D. glomerata* plants as the *E. typhina* strain infected those plants formed stromata on every infected tiller. RNA was extracted from all basal, sub-stromata (*D. glomerata* only), stromata and asymptomatic inflorescence (*Elymus* sp. only) samples (2.7.2), and cDNA was generated from these RNA samples (2.7.3). This cDNA was then diluted 5-fold and used as template for qPCR (2.6.7) using the primer pairs described in Table 6.3.

The translation elongation factor-encoding genes *EF2* and *tefA* and ribosome small subunit component-encoding genes *rps10* and *rps22* were selected as reference genes for this qPCR analysis based on a previous study that showed expression of those genes was stable across four different symbiosis defective *E. festucae* mutants (Chujo and Scott, 2014). Initial analysis of the *D. glomerata*/*E. typhina* results indicated that absolute expression levels of all reference genes were

approximately 2-fold reduced in fertilised vs. unfertilised stromata. Whole-tiller DNA was extracted (2.5.7) from the same material used to generate the RNA samples for qPCR to determine the reason for this apparent reduction in gene expression. Equal quantities of these DNA samples were used as templates for qPCR analysis (2.6.7) using primer pairs that amplify the *D. glomerata* 5.8S rRNA gene or the cytosolic heat shock protein 90.2 gene (Table 6.3) along with primers that amplify the *E. typhina rps22* and *tefA* genes to determine the approximate ratio of plant to fungal genomes present in these samples. The results showed that the number of fungal nuclei was increased approximately 40-fold in both fertilised and unfertilised stromata relative to the basal tissue samples. This increase is consistent with the massive increase in *Epichloë* hyphae associated with stroma formation, and demonstrated that the number of fungal nuclei did not decrease in fertilised vs. unfertilised stromata.

The expression of a perfect reference gene should directly correlate to the number of nuclei present. The increases in absolute gene expression of *EF2*, *tefA*, *rps10* and *rps22* in unfertilised stromata vs. basal tissue samples and fertilised stromata vs. basal tissue samples were therefore compared to the observed ~40-fold increase in fungal nuclei in all stromata vs. basal tissues. Absolute gene expression increases close to 40-fold in amplitude were identified for both *EF2* and *tefA* in unfertilised stromata vs. basal tissues, whereas the increase in expression of *rps10* and *rps22* expression across these same tissues was approx. 80-fold. In contrast, gene expression of *rps10* and *rps22* was increased approximately 40-fold in fertilised stromata vs. basal tissues, while *EF2* and *tefA* gene expression increased only ~20-fold. This indicated that increase in expression of *EF2* and *tefA* was equivalent to the increase in number of fungal nuclei for unfertilised stromata vs. basal tissues. However, the increase in *EF2* and *tefA* expression was only half of the observed increase in fungal nuclei population for fertilised stromata vs. basal tissues. Conversely, expression of *rps10* and *rps22* was equivalent to the increase in fungal nuceli in fertilised stromata vs basal tissues only, with an increase in expression of these genes approximately 2-fold higher than the observed increase in fungal nuclei for unfertilised stromata vs basal tissues. Subsequent analyses were therefore performed using *EF2* and *tefA* as the reference genes for unfertilised *E. typhina* and

E. elymi stromata samples, while *rps10* and *rps22* were used as the reference genes for fertilised *E. typhina* stromata samples.

The qPCR results revealed that expression of *irlA* cluster genes did not differ significantly between *D. gomerata*/*E. typhina* basal and sub-stromata samples (Figure 6.7). Expression of all five *irlA* cluster genes was significantly upregulated in unfertilised *E. typhina* stromata vs. vegetative (basal/sub-stroma) tissue samples, with gene expression fold-increases of ×14 (*pdtA*), ×5 (*afrA*), ×3 (*irlA*), ×5 (*mfsB*) and ×5 (*fxbA*) observed (Figure 6.7A). Expression of *pdtA*, *mfsB* and *fxbA* was also upregulated in fertilised *E. typhina* stromata relative to the vegetative tissues, but the amplitude of this increase was less than that observed for unfertilised stromata (Figure 6.7B). Expression levels of *afrA* and *irlA* in fertilised *E. typhina* stromata were not significantly different to the vegetative tissues (Figure 6.7B). The *irlA* cluster gene regulation patterns contrasted those observed for genes *EF320* and *esdC*, which were located immediately downstream of the *irlA* gene cluster. *EF320* gene expression was significantly downregulated in both fertilised and unfertilised stromata relative to vegetative tissues, while the levels of *esdC* mRNA did not change significantly between tissue types (Appendices 9.4.5 & 9.4.1). The patterns of regulation for the *irlA* cluster genes and *esdC* in *E. elymi* was similar to the patterns observed for those genes in *E. typhina* (Figure 6.8; Appendix 9.4.1).

Table 6.3 Primer pairs used for RT-qPCR analysis of *irlA* locus gene expression in *E. typhina* and *E. elymi*

Primer 1	Primer 2	Product size (bp)	Target
pdtA_3	pdtA_4	144	3' end of *pdtA* (*E. typhina* & *E. elymi*)
afrA_3	afrA_4	90	3' end of *afrA* (*E. typhina* & *E. elymi*)
irlA_3	irlA_4	114	3' end of *irlA* (*E. typhina* & *E. elymi*)
mfsB_3	mfsB_4	110	3' end of *mfsB* (*E. typhina* only)
mfsB_3	mfsb_4.1	110	3' end of *mfsB* (*E. elymi* only)
fxbA_3	fxbA_4	145	3' end of *fxbA* (*E. typhina* & *E. elymi*)
TC399	TC400	82	3' UTR of *EF2* (*E. typhina* only)
TC399.1	TC400	82	3' UTR of *EF2* (*E. elymi* only)
TC403.1	TC404.1	82	3' UTR of *rps10* (*E. typhina* & *E. elymi*)
TC407	TC408	83	3' UTR of *rps22* (*E. typhina* only)
TC407	TC408.1	83	3' UTR of *rps22* (*E. elymi* only)
tefA_1	tefA_2	137	Across 4th *tefA* intron (*E. typhina* & *E. elymi*)
esdC_1	esdC_2	134	3' end of *esdC* (*E. typhina* and *E. elymi*)
EF320_1	EF320_2	99	3' end of *EF320* (*E. typhina* only)
Dgl_1	Dgl_2	85	5.8S rRNA gene (*D. glomerata* only)
Dgl_5	Dgl_6	87	heat shock protein 90.2 gene (*D. glomerata* only)

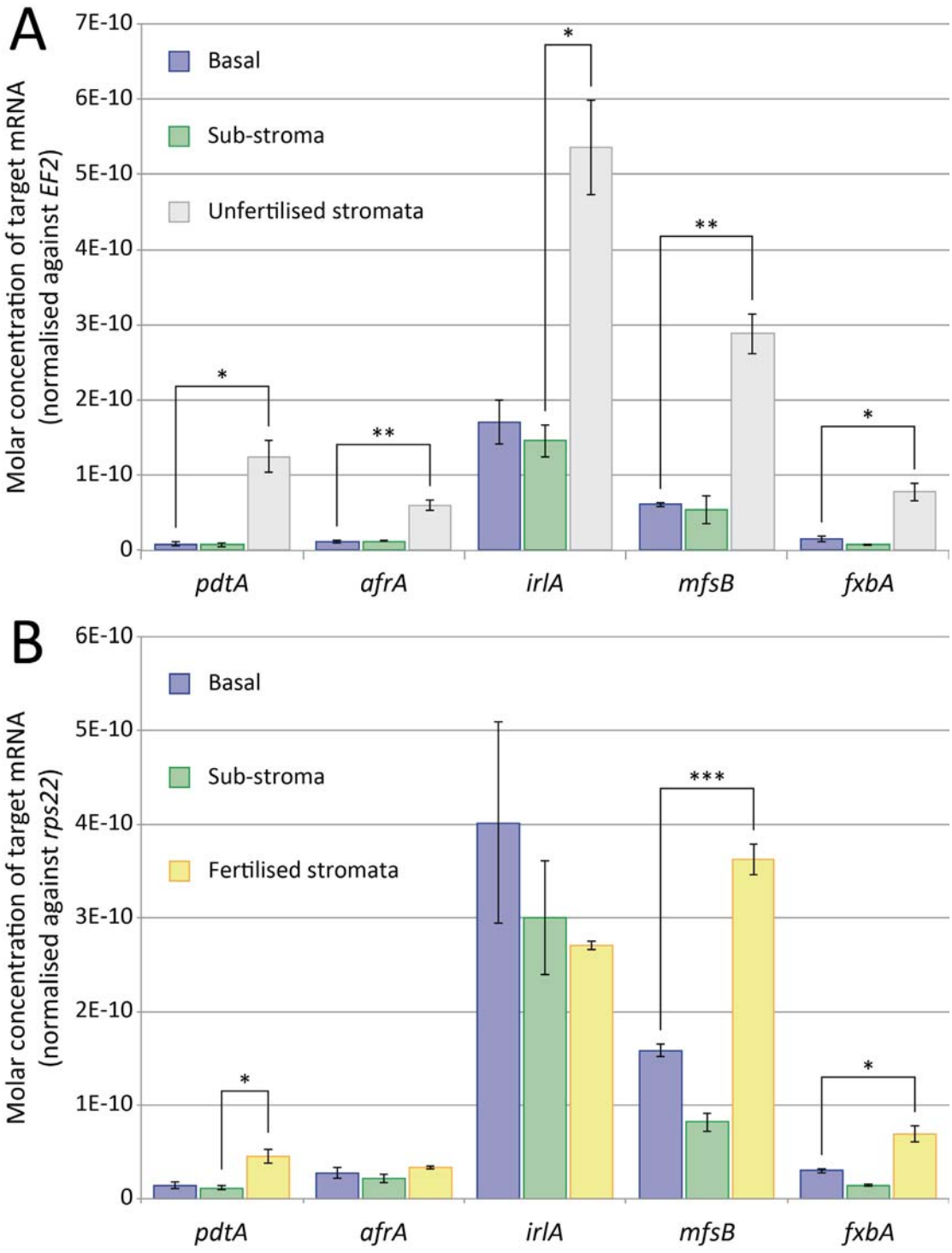


Figure 6.7 Expression of *E. typhina irlA* cluster genes is upregulated in stromata. (A) Expression of *irlA* cluster genes in unfertilised stromata compared to basal and sub-stroma tissues. (B) Expression of *irlA* cluster genes in fertilised stromata compared to basal and sub-stromata tissues. Error bars show the standard error of the mean. Significance of difference between means calculated using one-tailed Welch's *t*-test, with the tissue comparison exhibiting the weakest significance shown for each gene.

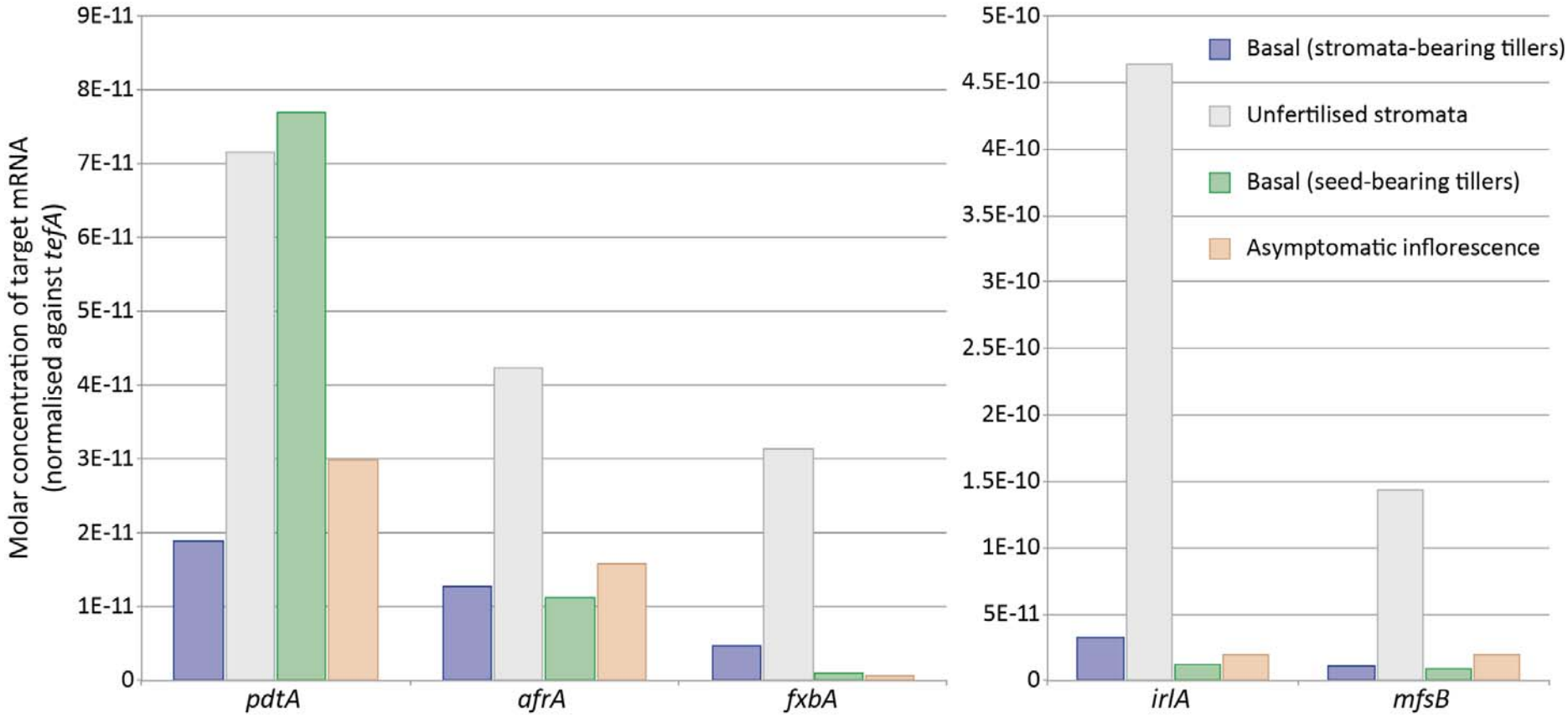


Figure 6.8 Expression of *E. elymi irlA* cluster genes appears to be upregulated in unfertilised stromata. Expression of *irlA* cluster genes was compared between basal tissues [from tillers bearing either stromata or seed (asymptomatic inflorescence)], unfertilised stromata and asymptomatic inflorescences. Gene expression was normalised against *tefA*. Statistical analyses not performed due to insufficient replicates.

6.7 Generation of individual *irlA* cluster gene deletion strains

Deletion constructs for the *irlA* cluster genes *pdtA*, *afrA*, *mfsB* and *fxbA* were excised using appropriate restriction enzymes from plasmids pDB24 (9.3.14), pDB25 (9.3.15), pDB26 (9.3.16) and pDB27 (9.3.17), respectively. These constructs were purified by gel extraction (2.6.9) then used in separate transformations of *E. festucae* Fl1 protoplasts (2.9.2). Colonies generated from the *pdtA*, *afrA*, *mfsB* and *fxbA* deletion construct transformants were then screened by PCR (200 colonies from each transformation; 2.6.3) using primer pairs 24SCR_3 / hph_iR, 25SCR_3 / hph_iF, 24SCR_4 / hph_iF and 27SCR_4 / hph_iR, respectively. These primer pairs were all designed to enable PCR amplification from within the P*trpC-hph* cassette across one border of the integrated deletion constructs, with amplification only expected for mutants in which a deletion construct had integrated by homologous recombination at the target locus. This PCR screen identified eight potential Δ*pdtA* mutants, five potential Δ*afrA* mutants, four potential Δ*mfsB* mutants and five potential Δ*fxbA* mutants. However, PCR amplification of gDNA template from each of these transformants with primer pairs 24SCR_5 / 24SCR_6, 25SCR_1 / 25SCR_2, 26SCR_5 / 26SCR_6 and 27SCR_5 / 27SCR_6, all of which were designed to amplify within the deleted region of the target genes, revealed that these potential gene deletion mutants were all heterokaryotic mixtures of targeted gene deletion strains and strains that only contained ectopic integrations of the deletion construct. Single spore isolation (2.4.2.4) was used to attempt purification of targeted integration mutant strains from these potential heterokaryons, with isolation achieved for strains Δ*pdtA*-168, Δ*afrA*-11, Δ*mfsB*-104, Δ*mfsB*-153, Δ*fxbA*-15 and Δ*fxbA*-137 (Figure 6.9). These gene deletion strains were subsequently screened by Southern analysis to determine the integration patterns of each construct and identify transformants that were the result of a single-copy integration (Figure 6.10; Figure 6.11; Figure 6.12; Figure 6.13). Analysis of those results revealed that strains Δ*mfsB*-153 and Δ*fxbA*-15 contained single-copy integrations of their respective deletion constructs at the target loci, strains Δ*pdtA*-168, Δ*afrA*-11 and Δ*mfsB*-104 contained integration of deletion construct concatamers at their respective target loci, and Δ*fxbA*-137 contained a complex integration pattern. Δ*pdtA*-168, Δ*afrA*-11 and

Δ*mfsB*-104 also contained additional ectopic integrations of their respective deletion constructs.

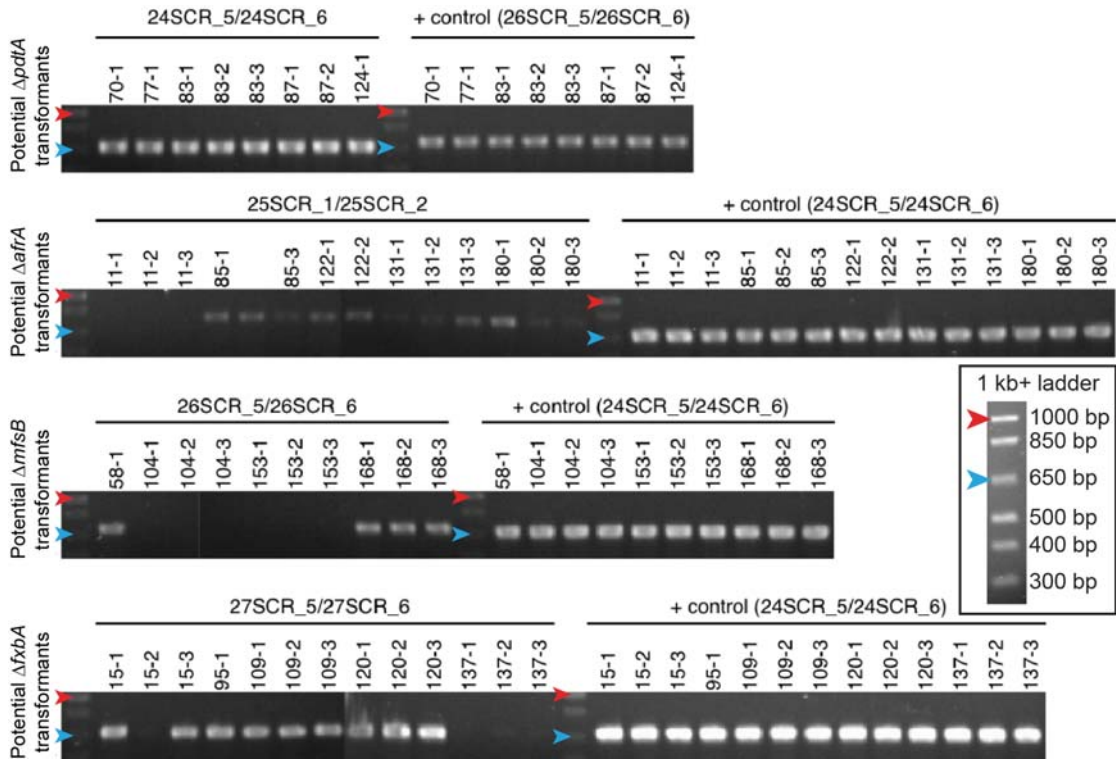


Figure 6.9 PCR screen of spore-derived strains. PCR screen of gDNA from single-spore isolated colonies using primer pairs 24SCR_5/24SCR_6 (668 bp expected product), 25SCR_1/25SCR_2 (785 bp expected product), 26SCR_5/26SCR_6 (704 bp expected product) and 27SCR_5/27SCR_6 (678 bp expected product), which amplify within *pdtA*, *afrA*, *mfsB* and *fxbA*, respectively. Primer pair 26SCR_5/26SCR_6 was used as a template integrity control for pDB24 transformants and primer pair 24SCR_5/24SCR_6 was used as a template integrity control with all other transformants. Templates that produced a product for the control PCR reaction and no product for the screening PCR reaction were likely candidates for homokaryotic targeted gene deletion strains. Inset: 1 kb+ ladder (Invitrogen). The 1000 bp and 650 bp ladder bands are indicated on the gel photos using red and blue arrowheads, respectively.

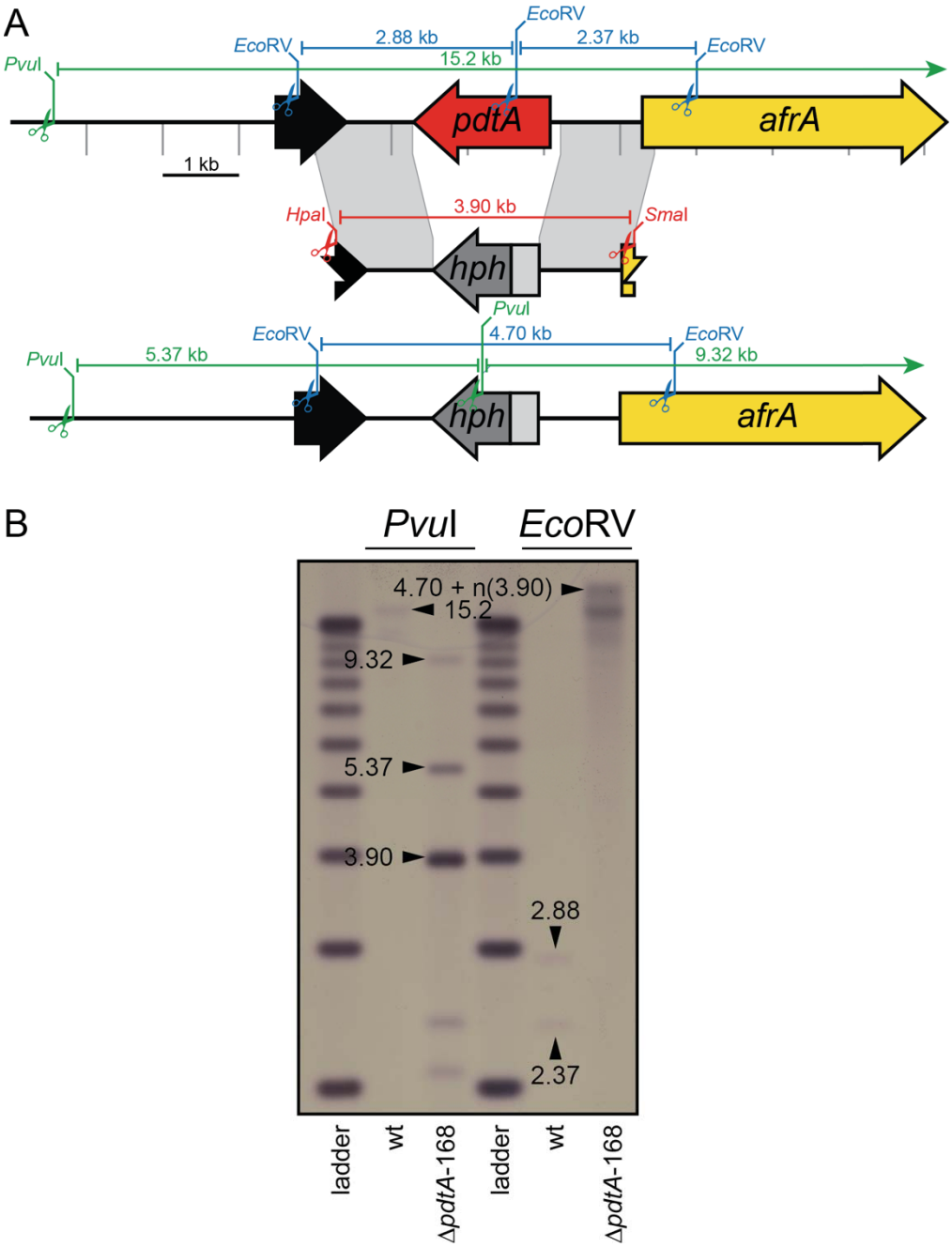


Figure 6.10 Southern blot of mutant Δ*pdtA*-168. (A) Top: gene map of wt *pdtA* locus annotated with relevant fragments expected from digestion with *Pvu*I or *Eco*RV. Middle: gene map of the *pdtA* deletion construct after excision from pDB24 using *Hpa*I and *Sma*I. Bottom: gene map of the anticipated Δ*pdtA* locus annotated with relevant fragments expected from digestion with *Pvu*I or *Eco*RV. (B) Southern blot of wt and Δ*pdtA*-168 gDNA digested with *Pvu*I or *Eco*RV and probed with DIG-dUTP labelled circular pDB24 plasmid. Band sizes of expected fragments are annotated. The ladder (1 kb plus; Invitrogen) hybridises to the pDB24-derived probe due to homologous DNA sequences shared between the pDB24 vector backbone and ladder fragments ≥ 1 kb in size. The smallest (lowest) ladder band in this image is 2 kb in size, and each subsequent band (ascending) is 1 kb larger than the previous band. The largest band of the ladder is 12 kb in size.

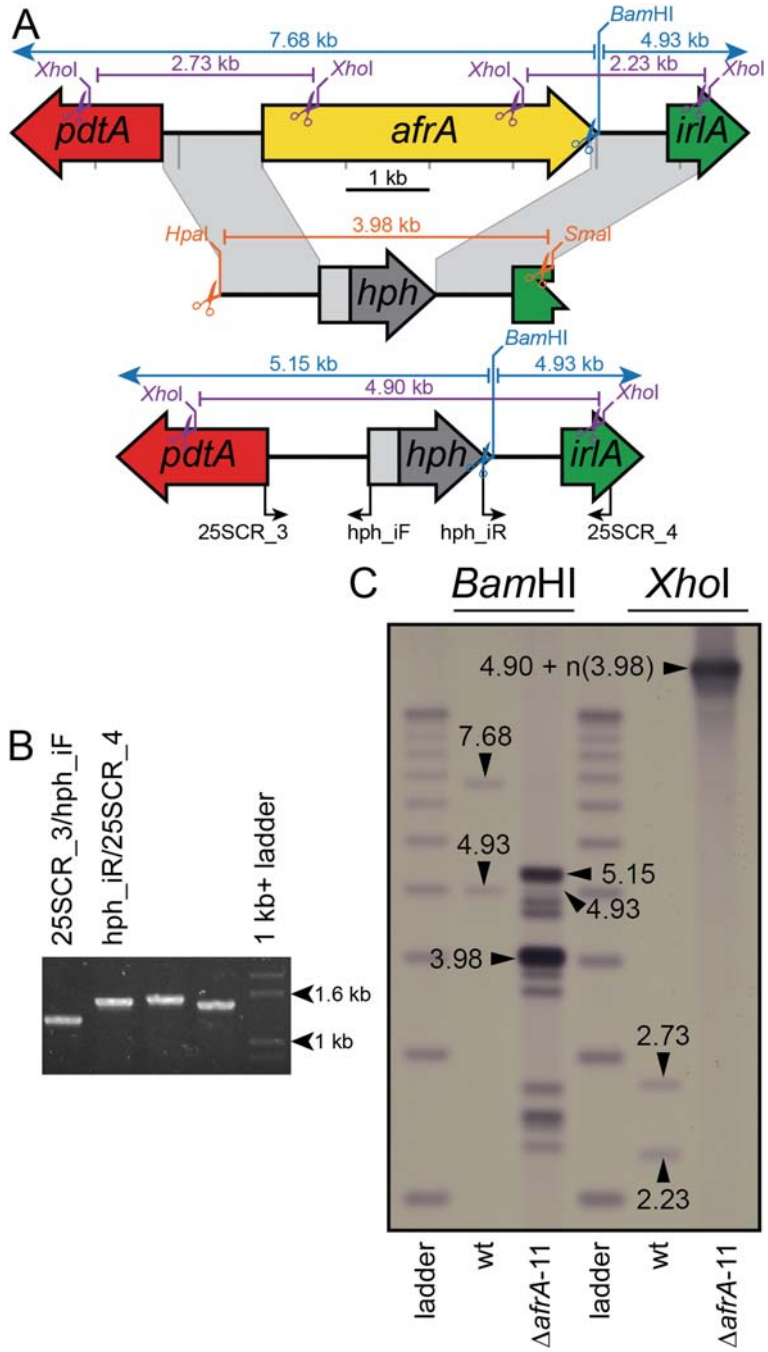


Figure 6.11 Southern blot of mutant Δ*afrA*-11. (A) Top: gene map of wt *afrA* locus annotated with relevant fragments expected from digestion with *Bam*HI or *Xho*I. Middle: gene map of the *afrA* deletion construct after excision from pDB25 using *Hpa*I and *Sma*I. Bottom: gene map of the anticipated Δ*afrA* locus annotated with relevant primers and fragments expected from digestion with *Bam*HI or *Xho*I. (B) PCR analysis of Δ*afrA*-11 gDNA using primer pairs 25SCR_3/hph_iF (expected product 1259 bp) and hph_iR/25SCR_4 (expected product 1525 bp), which amplify from within the P*trpC–hph* cassette across the left and right borders of the *afrA* deletion construct, respectively. Unlabelled bands are not relevant to this analysis. (C) Southern blot of wt and Δ*afrA*-11 gDNA digested with *Bam*HI or *Xho*I and probed with DIG-dUTP labelled circular pDB25 plasmid. Band sizes of expected fragments are annotated. Digests are complex due to the presence of one or more ectopic deletion construct insertions. The ladder (1 kb plus; Invitrogen) hybridises to the pDB25-derived probe due to homologous DNA sequences shared between the pDB25 vector backbone and ladder fragments ≥ 1 kb in size. The smallest (lowest) ladder band in this image is 2 kb in size, and each subsequent band (ascending) is 1 kb larger than the previous band. The largest band of the ladder is 12 kb in size.

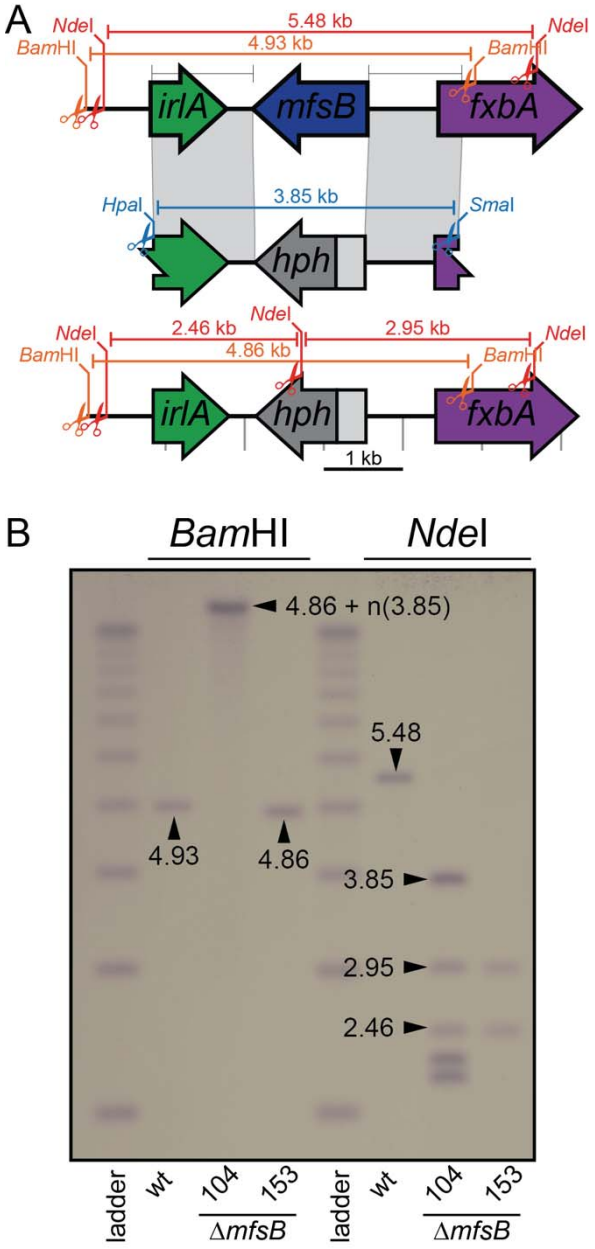


Figure 6.12 Southern blot of Δ*mfsB* mutants. (A) Top: gene map of wt *mfsB* locus annotated with relevant fragments expected from digestion with *Bam*HI or *Nde*I. Middle: gene map of the *mfsB* deletion construct after excision from pDB26 using *Hpa*I and *Sma*I. Bottom: gene map of the anticipated Δ*mfsB* locus annotated with 1 kb scale markers and relevant fragments expected from digestion with *Bam*HI or *Nde*I. (B) Southern blot of wt, Δ*mfsB*-104 and Δ*mfsB*-153 gDNA digested with *Bam*HI or *Nde*I and probed with DIG-dUTP labelled circular pDB26 plasmid. Band sizes of expected fragments are annotated. The ladder (1 kb plus; Invitrogen) hybridises to the pDB26-derived probe due to homologous DNA sequences shared between the pDB26 vector backbone and ladder fragments ≥ 1 kb in size. The smallest (lowest) ladder band in this image is 2 kb in size, and each subsequent band (ascending) is 1 kb larger than the previous band. The largest band of the ladder is 12 kb in size.

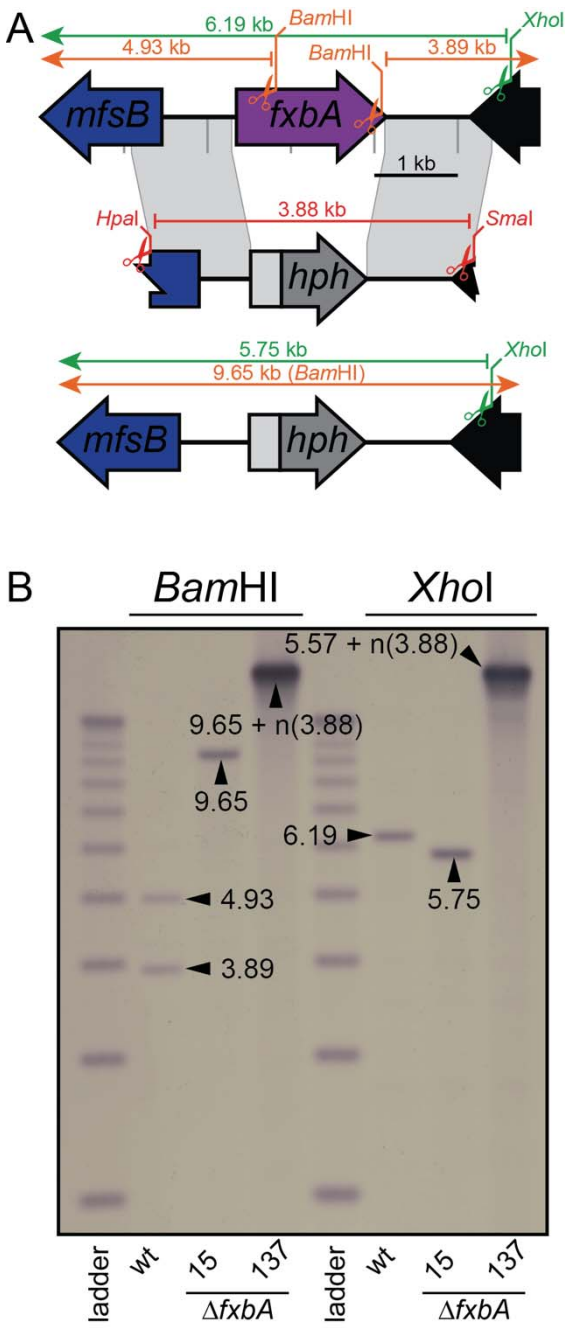


Figure 6.13 Southern blot of Δ*fxbA* mutants. (A) Top: gene map of wt *fxbA* locus annotated with relevant fragments expected from digestion with *Bam*HI or *Xho*I. Middle: gene map of the *fxbA* deletion construct after excision from pDB27 using *Hpa*I and *Sma*I. Bottom: gene map of the anticipated Δ*fxbA* locus annotated with relevant fragments expected from digestion with *Bam*HI or *Xho*I. (B) Southern blot of wt, Δ*fxbA*-15 and Δ*fxbA*-137 gDNA digested with *Bam*HI or *Xho*I and probed with DIG-dUTP labelled circular pDB27 plasmid. Band sizes of expected fragments are annotated. The ladder (1 kb plus; Invitrogen) hybridises to the pDB27-derived probe due to homologous DNA sequences shared between the pDB27 vector backbone and ladder fragments ≥ 1 kb in size. The smallest (lowest) ladder band in this image is 2 kb in size, and each subsequent band (ascending) is 1 kb larger than the previous band. The largest band of the ladder is 12 kb in size.

6.8 Phenotype of *irlA* cluster gene deletions

The phenotype of all *irlA* cluster gene deletion mutant cultures grown on solid media was the same as wt (data not shown). To investigate the phenotype of these mutants *in planta* strains Δ*pdtA*-168, Δ*afrA*-11, Δ*mfsB*-153 and Δ*fxbA*-137 were inoculated (2.10.1) along with wt and *ΔirlA*-62 controls (Δ*irlA*-487 had not been generated at this time) into 56 *Lolium perenne* seedlings each and grown under standard conditions (2.10.2) for 12 weeks post-inoculation (11 weeks post-inoculation for Δ*pdtA*-168). The length of each tiller on each surviving plant was then recorded, and infection status was determined by immunoblot (2.10.3). The results revealed that *L. perenne* plants infected with Δ*pdtA*, Δ*afrA*, Δ*mfsB* or Δ*fxbA* mutants exhibited the same macroscopic phenotype as wt during vegetative growth (Figure 6.14; Figure 6.15), sharply contrasting the stunted phenotype of Δ*irlA*-62-infected plants, but similar to the phenotype of Δ*irlA*-487 plants (Figure 5.10). Average tiller length for plants infected with Δ*pdtA*, Δ*afrA*, Δ*mfsB* or Δ*fxbA* strains also did not differ significantly from wt-infected plants (Figure 6.16). Microscopic analysis of the endophyte *in planta* also did not reveal any phenotypic differences between plants infected with wt or any of the Δ*pdtA*, Δ*afrA*, Δ*mfsB* or Δ*fxbA* mutants (Figure 6.17). This was perhaps not surprising given that expression of these genes was very low in plant vegetative tissues (Figure 5.15; Figure 6.7). This also indicated that the mechanisms causing host stunting in the Δ*irlA*-62 strain were not induced by these *pdtA*, *afrA*, *mfsB* or *fxbA* gene deletions.

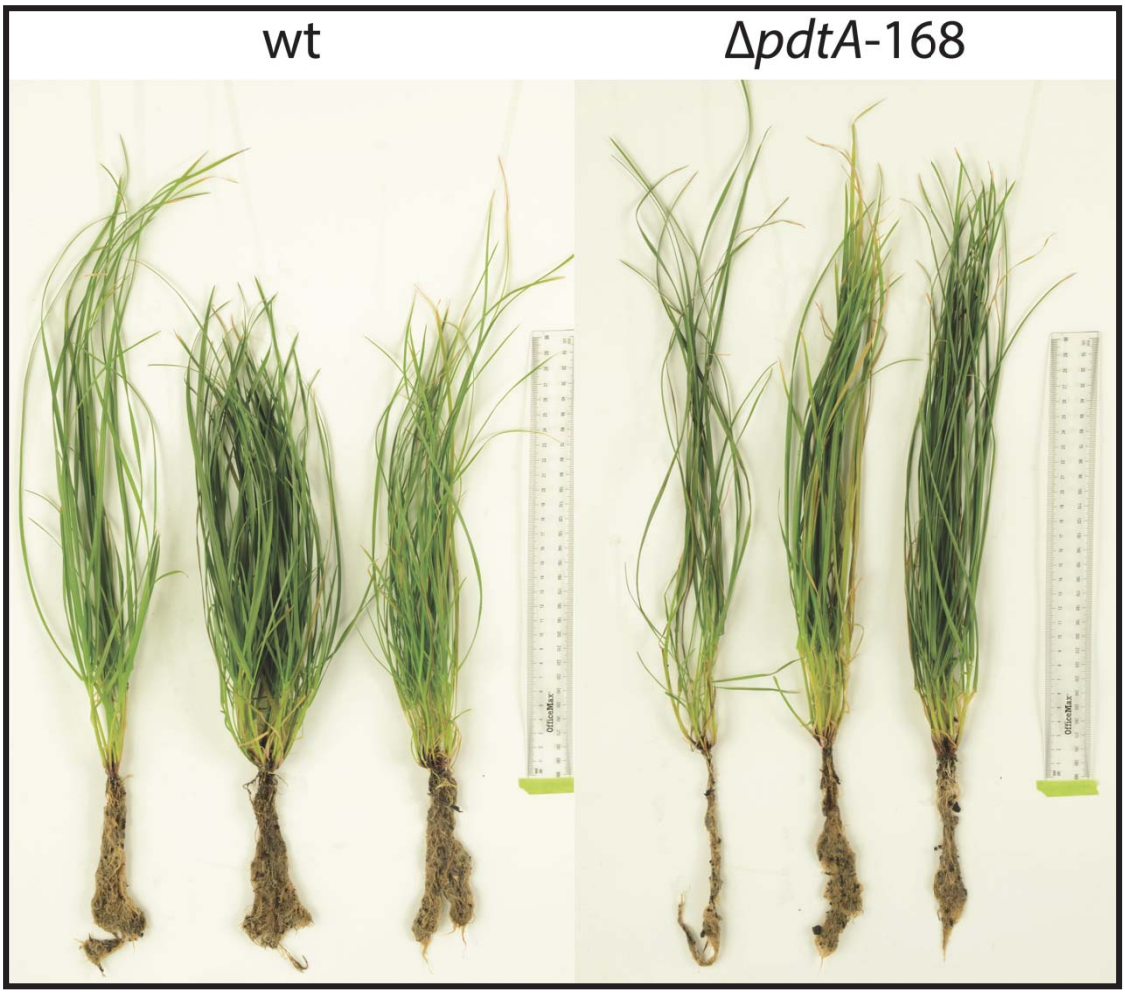


Figure 6.14 Phenotype of Δ*pdtA*-168-infected plants. Images show representative photos of wt and Δ*pdtA*-168 infected *L. perenne* plants at 11 weeks post-inoculation. 30 cm ruler included for scale.

.

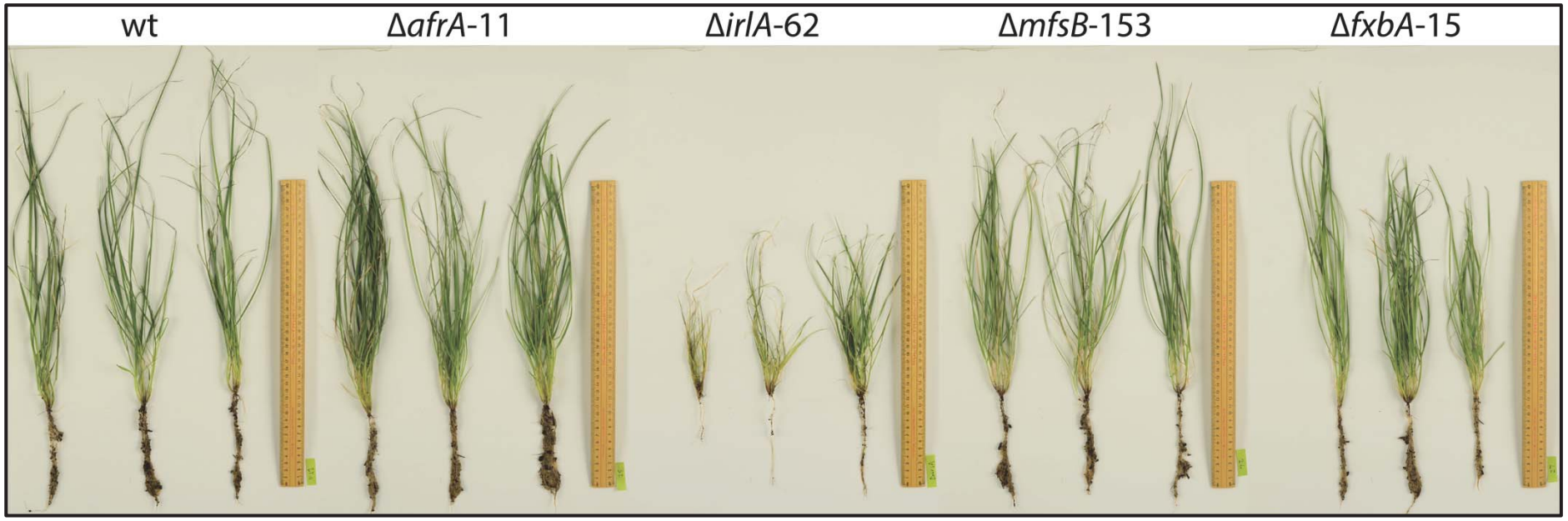


Figure 6.15 Phenotypes of plants infected with *irlA* cluster mutants. Images show representative *L. perenne* plants infected with Δ*afrA*-11, Δ*mfsB*-153 or Δ*fxbA*-15, Δ*irlA*-62 or wt at 12 weeks post-inoculation. 40 cm ruler included for scale.

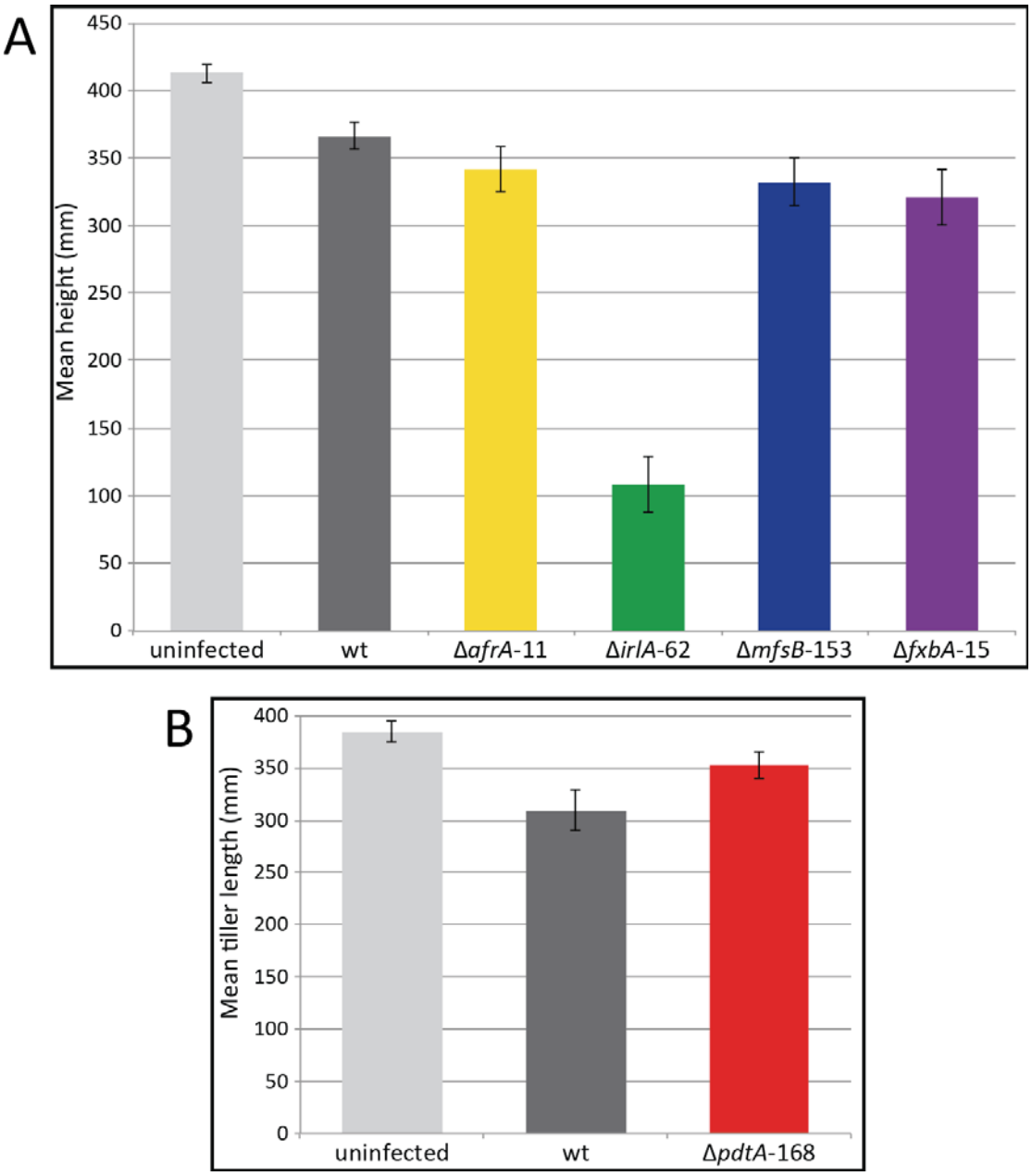


Figure 6.16 Mean tiller length of plants infected with *irlA* cluster gene deletion mutants. (A) Graph showing mean tiller length of 130 uninfected, 17 wt-infected, 16 Δ*afrA*-11-infected, 5 Δ*irlA*-62-infected, 22 Δ*mfsB*-153-infected and 17 Δ*fxbA*-15-infected *L. perenne* plants at 12 weeks post-inoculation. (B) Graph showing mean tiller length of 51 uninfected, 7 wt-infected and 20 Δ*pdtA*-168-infected *L. perenne* plants at 11 weeks post-inoculation. Data for (A) and (B) were generated from separate experiments. Error bars show the standard error of the mean.

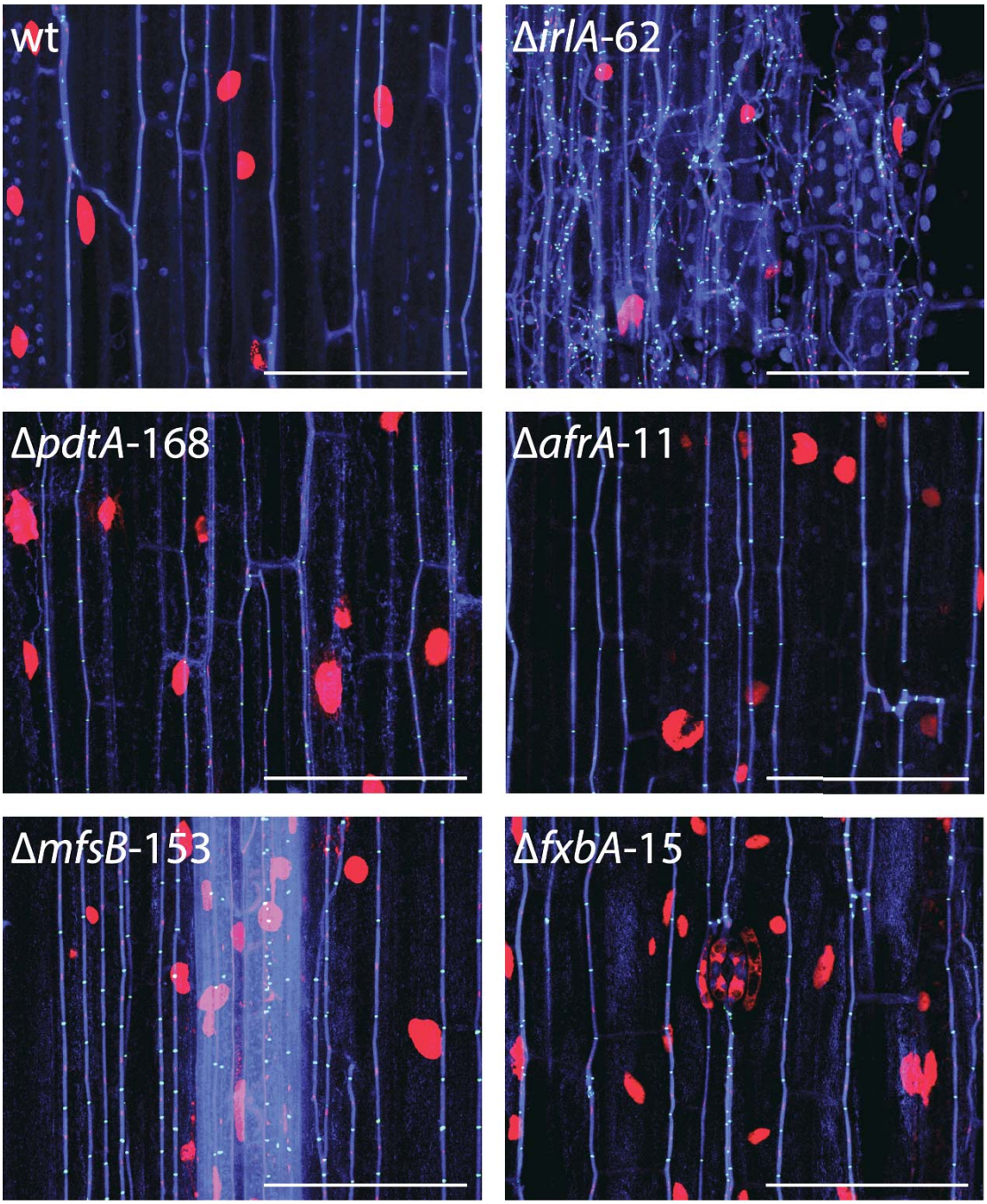


Figure 6.17 Microscopic phenotype of *irlA* cluster deletion mutants *in planta*. Images show CLSM images of endophytic hyphae growing just below the pseudostem surface. Images show 10 µm-thick stacks with 0.5 µm separation between slices. Green shows the location fungal chitin, mainly found in the fungal septa. Red shows both fungal and plant nuclei. Blue shows the fungal cytoplasm and some plant tissues, particularly in vascular bundles. Scale bar length is 100 µm.

6.9 Discussion

The results presented in this chapter establish *irlA* as part of a five-gene cluster containing *pdtA*, *afrA*, *mfsB* and *fxbA*. These genes are predicted to encode the PLP-dependent transferase (PDT) PdtA, the adenylate-forming reductase (AFR) AfrA, the isoflavone reductase-like (IRL) protein IrlA, the major facilitator superfamily (MFS) transporter MfsB and the FAD-binding oxidase FxbA. These functions are consistent with a SM gene cluster. Evidence for this cluster is provided by synteny analyses across the Clavicipitaceae showing that the *irlA* cluster genes, when present, are always located contiguously. Genes belonging to the same biosynthetic pathway in fungi are commonly assembled into a contiguous metabolic gene cluster (Keller and Hohn, 1997; Keller et al., 2005). The observation that the *irlA* cluster genes follow this trend therefore suggests that *pdtA*, *afrA*, *irlA*, *mfsB* and *fxbA* likely encode components of the same biosynthetic pathway. The fact that the *irlA* gene clusters of *Epichloë* spp., *A. take*, *G. lozoyensis* and *Metarhizium* spp. were all located at different genomic loci supports this hypothesis, as it suggests that clustering of these genes was conserved after several independent relocation events. Furthermore, the high conservation of gene order and orientation between these homologous clusters indicates these events occurred through relocation of the entire *irlA* cluster in one event, as is typical for fungal SM gene clusters (Wisecaver and Rokas, 2015; Wisecaver et al., 2014).

A series of gene deletion and relocation events following divergence from a common ancestor can easily explain this *irlA* cluster distribution within the Clavicipitaceae. Discontinuous distribution of SM gene clusters, such as is observed for the *irlA* cluster across the Clavicipitaceae, is the norm for filamentous fungi (Keller et al., 2005; Wisecaver et al., 2014). *Epichloë* spp. are no exception to this trend, with discontinuous distributions described for a variety of *Epichloë* SM gene clusters within the literature (Berry et al., 2015b; Schardl et al., 2012; Schardl et al., 2013a; Young et al., 2009). This sharply contrasts the continuous distribution of the *irlA* gene cluster across sexual *Epichloë* spp., suggesting this cluster plays an important or even indispensable role in sexual *Epichloë* isolates. The presence of an *irlA* gene cluster in *G. lozoyensis* is surprising given the genetic distance separating

this species from the Clavicipitaceae. *G. lozoyensis* belongs to the Leotiomycetes (Youssar et al., 2012), whereas the Clavicipitaceae belong to the Sordariomycetes. It is therefore unusual that *G. lozoyensis irlA* cluster genes appear to be no more distantly related to their *Epichloë* homologs than the *Epichloë irlA* cluster genes are to their *A. take* homologs, as *A. take* is much more closely related to *Epichloë* spp. than *G. lozoyensis* is. These results may suggest a horizontal gene transfer event may have taken place between ancestors of *Epichloë* spp. and *G. lozoyensis*. The presence of the *irlA* cluster across multiple Clavicipitaceae genera suggests that the donor in such a transfer event would have belonged to the Clavicipitaceae. However, making a conclusive statement on the direction of transfer is complicated by the fact that *G. lozoyensis* is the only member of the family Helotiaceae for which the genome sequence has been made available to date. As further Helotiaceae genomes become available this question of transfer direction could be readdressed.

The gene composition of the *irlA* gene cluster is similar to that of the characterised *lna* and *lnb* gene clusters in *Aspergillus flavus* (Forseth et al., 2013). Like the *irlA* cluster, the *lna*/*lnb* cluster genes encode AFRs, IRL proteins, oxidases and MFS transporters (Figure 6.18). Forseth et al. (2013) show that the AFRs LnaA and LnaB encoded within these two clusters both reduce tyrosine, with the homologous genes *lnaA* and *lnbA* that encode these proteins apparently derived from a gene duplication event, sharing 60% nucleotide sequence identity. Interestingly, a Δ*lnaA*/KD::*lnbA* double mutant (KD = knock-down by RNAi), but not the Δ*lnaA* mutation alone, was shown to supress sclerotia formation in *A. flavus* (Forseth et al., 2013). Metabolite analyses of these mutants also showed that LnaA and LnbA can produce the same diastereomeric piperazine products (Forseth et al., 2013). Together, these data are proposed to demonstrate functional redundancy between these AFRs, though the authors do not address the possibility that the sclerotia suppression phenotype is due to the KD::*lnbA* alone. Forseth et al. (2013) also provide evidence that the IRL protein LnaB reduces the product produced by LnaA, and that expression of the encoding gene *lnaB* is upregulated by addition of the LnaA/LnaB pathway end product (molecule **1** in Figure 6.18) to the culture media. Forseth et al. (2013) also identified oxidised derivatives of the LnaA/LnbA products, and hypothesise that the oxidase-encoding genes within the *lna*/*lnb*

clusters catalyse the oxidation steps within this biosynthetic pathway. It is not clear if the the *lna* and *lnb* biosynthetic pathways remain similar, or diverge after the steps catalysed by LnaA/LnbA.

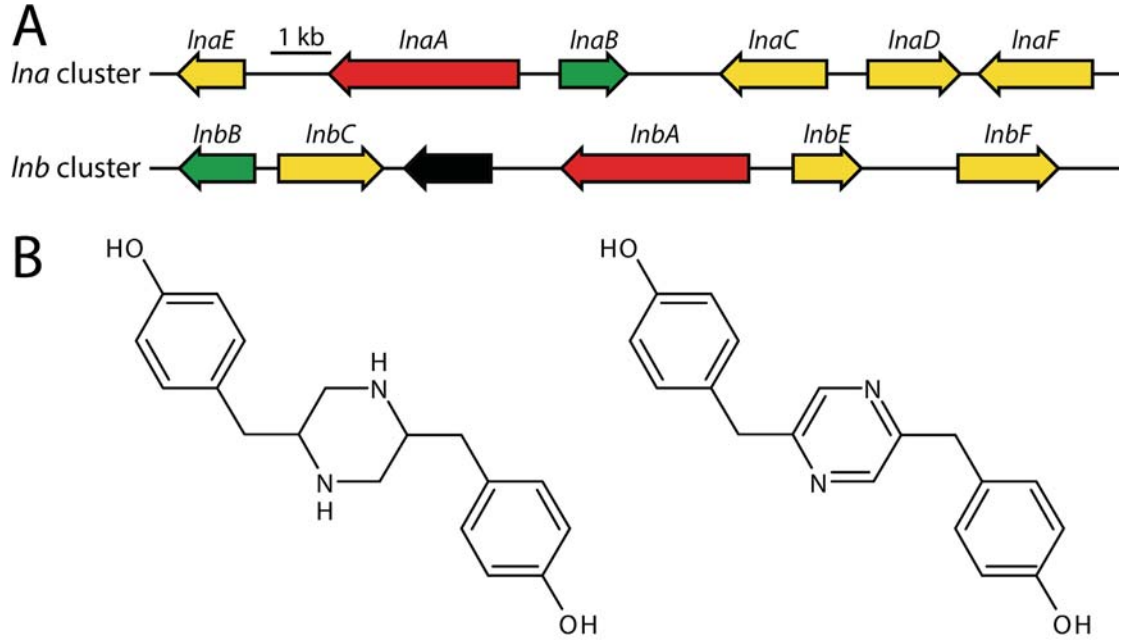


Figure 6.18 Locus map and products of the *lna* and *lnb* clusters of *Aspergillus flavus* as reported by Forseth et al., (2013). (A) Gene maps of the *lna* and *lnb* loci in *A. flavus*. The relationship between gene name suffix and encoded protein function is as follows: A = AFR, B = IRL protein, C & D = P450 oxidase, E = reductase and F = transporter. The genes *lnaA*/*lnbA* and *lnaB*/*lnbB* are coloured the same to indicate paralogy, while other cluster genes are coloured yellow. (B) End products of the *lna*/*lnb* clusters.

Interestingly, while *lnaA*/*lnbA* and *lnaB*/*lnbB* appear to be paralogs, perhaps arising from the same gene duplication event, comparison of the other genes proposed to comprise the *lna* and *lnb* gene clusters does not indicate significant sequence homology (<20% protein sequence identity). Despite this lack of homology, the proteins encoded by these genes belong to the same or similar gene families. This suggests that these genes have been independently recruited into these clusters after the proposed gene duplication event that gave rise to *lnaA*/*lnaB* and *lnbA*/*lnbB*. Given the evidence that the *lna* and *lnb* clusters encode at least

partially redundant pathways (Forseth et al., 2013), the possibility that these non-homologous genes recruited by each cluster may perform the same or similar functions has intriguing implications for the *irlA* cluster in *Epichloë* spp. AfrA and IrlA are not well conserved with LnaA and LnaB, sharing only 18% and 17% protein sequence identity, respectively, but they do belong to the same protein families. Like LnaA/LnbA, AfrA is also predicted to bind an aromatic amino acid substrate. Given the overall similarities in gene composition between the *irlA* and *lna*/*lnb* gene clusters, it is possible that the biosynthetic pathway driven by the *irlA* cluster gene products is similar to that of the *lna*/*lnb* clusters. Interestingly, a phenolic hydroxyl group is present in the substrates of all characterised PIP/IFR family proteins (Louie et al., 2007; Min et al., 2003; Wang et al., 2006), and has been shown to be required for substrate reduction by eugenol synthase (Koeduka et al., 2006). As IrlA appears to be closely related to PIP/IRL family proteins, this suggests that tyrosine may be the substrate of AfrA.

A number of oxidases are encoded by the *lna*/*lnb* gene clusters in *A. flavus*, including a putative alchohol dehydrogenase, NAD:flavin oxidoreductase and several cytochrome P450 monoxygenases (Forseth et al., 2013). The domain architecture of FxbA, consisting of a type-2 FAD-binding domain and C-terminal berberine bridge enzyme-like (BBE) domain, is not similar to any of these *lna*/*lnb* oxidases. One characterised protein for which the domain architecture closely matches FxbA is AfoF from *Aspergillus nidulans*. AfoF contains FAD-binding and BBE domains in the same arrangement as FxbA, though it contains a predicted signal peptide at its N-terminus instead of the single transmembrane domain observed in FxbA. Chiang et al. (2009) propose that AfoF oxygenates the α-diketone proton in one of the final steps during asperfuranone biosynthesis. The chanoclavine-I synthase EasE, which is an oxygenase involved in ergot alkaloid biosynthesis within the Clavicipitaceae, also exhibits this type-2 FAD-dependent oxidase/C-terminal BBE-like domain architecture (Young et al., 2015). Protein sequence identity between FxbA and AfoF/EasE is low (21% and 22% respectively), so functional assignments based on these proteins are highly speculative. However, AfoF/EasE-like oxygenations or *lna*/*lnb*-like ring oxidations represent possible candidate reactions.

The *lna*/*lnb* gene clusters of *A. flavus* do not contain a PLP-dependent transferase (PDT). However, a two-gene atromentin biosynthesis cluster has been previously characterised in the homobasidiomycete *Tapinella panuoides* that contains both the quinone synthase AtrA and the PDT AtrD (Schneider et al., 2008). The domain architecture of AtrA is very similar to that of the AFR proteins, containing an N-terminal A-domain followed by a PCP-domain. However, the C-terminal R-domain found in AFR proteins is replaced in AtrA by a TE-domain (Kalb et al., 2014; Schneider et al., 2008). The first step in this biosynthetic pathway is the deamination of L-tyrosine by AtrD, generating 4-hydroxyphenylpyruvic acid. AtrA then catalyses condensation of two 4-hydroxyphenylpyruvic acid molecules to form the final product, atromentin (Schneider et al., 2008). Unlike AtrA and AtrD, it is not a possibility that PdtA generates the substrate of AfrA through α-amino group deamination of a target amino acid. This can be determined by analysis of the AfrA A-domain 10AA substrate-binding code (DIHDIAAICK). The aspartic acid residue at position 1 and the lysine residue at position 10 are invariant across amino acid-binding A-domains. This is because the negatively charged sidechain of the asp residue is required to stabilise the positively charged α-amino group of the substrate amino acid, while the positively charged sidechain of the lys residue stabilises the substrate's negatively charged carboxylate group (Conti et al., 1997; May et al., 2002). In contrast, position 1 is occupied by a valine residue in A-domains that bind α-keto acids (Kalb et al., 2013). If PdtA is an α-amino group deaminase, its action must therefore occur downstream of AfrA in the *irlA* cluster biosynthetic pathway, as the substrate of AfrA is almost certainly an amino acid. This is by no means the only possible reaction that PdtA might catalyse, as PDT enzymes are known to facilitate an incredibly diverse range of amino acid-modifying reactions (di Salvo et al., 2013). It is important that any future experiments that express and purify the AfrA A-domain for substrate or structural determination include the 250 aa N-terminal extension of this protein, as it is likely a requirement for AfrA A-domain function (Kalb et al., 2015).

The observation that sclerotia formation is suppressed in the *A. flavus* Δ*lnaA*/KD::*lnbA* mutant has interesting implications for *irlA* cluster functionality. While *Epichloë* spp. are not known to produce sclerotia, the formation of stromata

as part of the *Epichloë* sexual cycle is well established (Clay and Schardl, 2002). Despite their different functions, sclerotia and stromata are structurally very similar. Both tissues are formed by pseudoparenchyma, but stromata bear fruiting bodies whereas sclerotia do not. The similarities between the *irlA* and *lna*/*lnb* clusters therefore suggest that the product(s) produced by the *irlA* cluster may play a role in stroma development. This hypothesis is supported by the observation that *irlA* cluster genes are highly conserved in the genomes of sexual *Epichloë* isolates, but not in asexual *Epichloë* isolates. Furthermore, the fact that upregulation of *pdtA*, *afrA*, *mfsB* and *fxbA* above basal levels was only observed in stromata supports this hypothesis. This stromata-specific expression pattern has now been confirmed across a diverse range of plant-endophyte associations, including *E. festucae*/*F. pratensis*, *E. festucae*/*L. perenne*, *E. elymi*/*Elymus* sp. and *E. typhina*/*Dactylis glomerata*, suggesting the this cluster is important for stroma development across all sexual *Epichloë* spp. Interestingly, in *Epichloë* spp. the *irlA* cluster is located immediately upstream of the genes *EF320* and *esdC*. These two genes are known targets of the transcription factor ProA, which is thought to be a repressor of sexual development in *Epichloë* spp. (Tanaka et al., 2013). Deletion of *proA* in *E. festucae* also results in dramatic downregulation of *irlA in planta* (Eaton et al., 2015), although ProA recognition sequences could not be identified in the promoter of *irlA*.

The fact that stroma development requires the endophyte to switch from restricted, mutualistic growth to proliferative growth suggests that the *irlA* cluster may function as a promoter of hyphal proliferation. This hypothesis is based on the results presented in Chapter 5, where it is shown that the pDB02-derived *irlA* gene deletion strains, which are proposed to also knockdown *mfsB* expression (KD::*mfsB*), induce a symbiosis-defective phenotype *in planta*. It is also shown that *irlA* is the only gene expressed at significant levels during vegetative growth *in planta*, and that upregulation of *pdtA*, *afrA* and *fxbA in planta* is associated with the Δ*irlA*/KD::*mfsB* symbiosis-defective phenotype. This suggests that *irlA* acts as a safeguard, ensuring that any intermediate product from the *irlA* cluster biosynthetic pathway is converted into a form that does not induce hyphal proliferation. Based on the results presented here and comparison to fungal SM pathways catalysed by similar proteins (Chiang et al., 2009; Forseth et al., 2013; Schneider et al., 2008), it is predicted that

the proteins encoded within the *irlA* gene cluster catalyse a branched biosynthetic pathway based on an aromatic amino acid substrate (Figure 6.19). The first committed step in this pathway is likely catalysed by AfrA, which would selectively bind and reduce an aromatic amino acid substrate, converting the substrate α-carboxylic acid to an aldehyde or alcohol. AfrA might also catalyse the formation of a non-peptide dimer from two reduced substrate molecules through a tethered imine intermediate, similar to the mechanism proposed for LnaA and LnbA in *A. flavus* (Forseth et al., 2013). PdtA would then catalyse the addition, removal or replacement of a specific functional group(s) on this intermediate, though it is also possible that PdtA would modify the sidechain of the substrate amino acid prior to reduction by AfrA, or after oxidation by FxbA. Finally, reduction by IrlA and/or oxidation by FxbA are predicted to result in two possible end products, with the FxbA product resulting in hyphal proliferation. MfsB might be a transporter for the proliferation-promoting product. This branched-pathway model would also explain why the symbiosis-defective phenotype associated with the Δ*irlA*/KD::*mfsB* mutants was not observed for other *irlA* cluster gene deletion mutants, as these would be predicted to be deficient in production of the proliferation-promoting end product.

The hypothetical proliferation-promoting product of the *irlA* cluster proteins may have to reach a certain threshold before hyphal proliferation and subsequent stroma development begins. This suggests a mechanism that could explain why sexual *Epichloë* species and strains differ dramatically in the frequency with which they form stromata over emerging host inflorescences. The relative balance of IrlA to the other pathway proteins, and perhaps FxbA in particular, is predicted to control the stromata formation frequency. Increasing the FxbA:IrlA ratio would be expected to increase the frequency with which stromata are formed. Based on the model presented here I predict that single-copy deletion construct integration Δ*irlA* mutations, which are asymptomatic during vegetative growth, would cause a dramatic increase in stromata formation frequency during host flowering. Conversely, Δ*afrA*, Δ*fxbA*, Δ*pdtA* and perhaps also Δ*mfsB* gene deletion strains, which were all asymptomatic during vegetative growth, would be expected to abolish, inhibit or impair stromata formation. The hypothesis that *irlA* cluster gene deletion mutants alter host stromata formation frequency could be tested by

comparing the occurance of stromata on plants infected with these mutants to wt plants after vernalisation. Interestingly, this may also mean that naturally occurring mutations that abolish *afrA*, *fxbA*, *pdtA* and perhaps *mfsB* function could result in that isolate becoming asexual. In fact, such mutations were observed in several asexual non-hybrid strains (Figure 6.6). As asexual strains are desirable for agricultural use (Johnson et al., 2013a; Young et al., 2013), this suggests that the results presented in this chapter may find agricultural applications in the future.

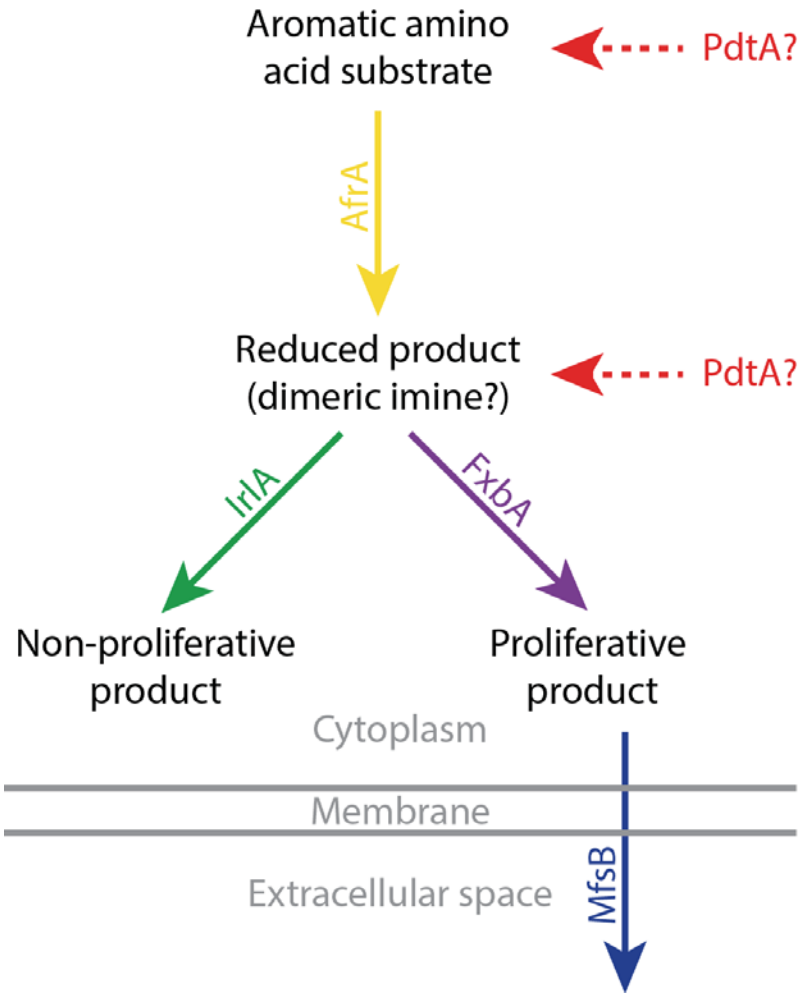


Figure 6.19 Speculative biosynthetic pathway for the *irlA* cluster gene products. IrlA might also act sequentially following modification of the reduced product by FxbA, or vice versa.

7 Conclusions

7.1 Conclusions: Chapters 3 & 4

Chapters 3 and 4 focused on the *Epichloë* SM gene *perA*, which encodes a two-module NRPS that produces the insect deterrent peramine. The first main area of research was a survey of *perA* alleles across *Epichloë* to identify mutations responsible for abolishing peramine production. This in turn enabled the design of rapid PCR and sequence-based screens to predict the peramine chemotype of novel isolates. Finally, the peramine-negative *perA-ΔR* allele was functionally characterised with the aim of determining whether the encoded protein was capable of producing a novel dipeptide product. The conclusions drawn based on this research are described below.

The discontinuous distribution of peramine production across *Epichloë* is primarily the result of small frameshift-causing indels and larger deletions within the *perA* gene. Complete loss of the *perA* gene was confined to the *E. glyceriae* and *E. gansuensis*/*E. inebrians* outgroups, though several large deletions within the *perA* gene were identified in some isolates. The distribution of specific *perA* null alleles was always confined within a single *Epichloë* spp., with the exception of the *perA-ΔR* allele.

The genetic basis of peramine-negative chemotypes was determined by PCR and/or *perA* gene sequence analysis for all isolates tested. A multi-tier screen to predict the peramine chemotype of novel isolates was designed based on the surveyed *perA* gene sequences to identify large deletions by PCR. Sequencing *perA* with conserved primers to identify smaller null allele-causing mutations followed. This method was also applicable to hybrid *Epichloë* spp. through bespoke sequencing primer design to polymorphic *perA* regions identified from an initial round of *perA* sequencing using the conserved *perA* sequencing primers. There were no instances in this study where the peramine negative chemotype of an isolate could not be explained genetically after PCR and/or *perA* sequence analysis.

The *perA-ΔR* allele was the only peramine-negative allele identified across multiple *Epichloë* spp. The *perA-ΔR* allele was identified within the sexual species *E. baconii*, *E. bromicola*, *E. festucae*, *E. sylvatica*, *E. typhina*, *E. typhina* ssp. *clarkii* and *E. typhina* ssp. *poae*, as well as the asexual species *E. siegelii* and *E.* sp. e4768. The phylogeny of *E. baconii*, *E. bromicola* and *E. festucae*-derived *perA-ΔR* alleles was also shown to exhibit apparent trans-species polymorphism. Evidence was provided suggesting that the *E. baconii*, *E. bromicola* and *E. festucae perA-ΔR* alleles are the result of recombination between ancestral *perA* and *perA-ΔR* alleles.

***Penicillium paxilli* is a suitable heterologous expression host for *perA*.** Expression of *perA* under the control of the *paxM* promoter successfully established peramine production in *P. paxilli*. This result strongly supports the hypothesis of Tanaka et al. (2005) that *perA* is the only pathway-specific gene required to enable peramine production.

PerA-ΔR proteins contain unique amino acid substitutions that alter substrate specificity. The PerA-ΔR-unique substitutions within the A1-domain substrate-binding region of PerA-ΔR proteins from *E. festucae* E2368 and *E. typhina* ssp. *poae* E1022 were shown to significantly reduce peramine production when introduced into PerA.

Replacement of the PerA-ΔR methylation domain is required to restore peramine production to *E. festucae* E2368 PerA-ΔR. In addition to the PerA A1-domain substrate-binding region and R-domain, replacement of the PerA-ΔR M-domain with its PerA counterpart was required to restore peramine production to *E. festucae* E2368 PerA-ΔR. This indicates that any PerA-ΔR product is likely to contain an arginine moiety with an altered methylation state relative to peramine. Replacement of the PerA-ΔR C-domain was not essential.

PerA-ΔR proteins appear likely to remain functional. Unfortunately, due to equipment failures, the PerA-ΔR protein has not yet been analysed for production of a novel product, and the release mechanism has not yet been analysed.

Nevertheless, the high conservation of *perA-ΔR* alleles, PerA-ΔR-unique A1-domain substrate-binding residue substitutions and apparent maintenance of TSP—which requires balancing selection—for *perA-ΔR* alleles suggests that PerA-ΔR proteins are likely to produce a product that provides a selective advantage to the endophyte.

7.2 Conclusions: Chapters 5 & 6

Chapters 5 and 6 focused on characterisation of the symbiotically regulated gene *irlA* and the putative five-gene cluster within which it is located. Initial research included generation and analysis of an *irlA* gene deletion mutant, with a focus on how that genotype affected the *E. festucae–L. perenne* symbiosis. Subsequent work established *irlA* as a member of a putative five-gene SM cluster. Bioinformatic, gene deletion and RT-qPCR analysis of these *irlA* cluster genes was utilised to gain insight into potential functions for this cluster. The conclusions drawn based on this research are described below.

Deletion of the *irlA* gene alone does not induce a symbiotic phenotype. The *E. festucae* Δ*irlA*-487 strain, which contains a single copy of the *irlA* deletion construct integrated at the target locus, did not induce a symbiotic phenotype. However, the pDB02-derived Δ*irlA* strains were shown to induce a strong symbiotic phenotype. Evidence was provided suggesting that this symbiotic phenotype is the result of both the *irlA* gene deletion and RNAi-induced post-transcriptional gene silencing of the downstream gene *mfsB*.

***irlA* is part of a putative five-gene secondary metabolism cluster.** Synteny analysis of the *irlA* locus both within *Epichloë* and across other Clavicipitaceaen genomes identified a putative five-gene cluster within which *irlA* is located. Based on the predicted functions of the proteins encoded within this cluster, a SM role is proposed.

Deletion of *irlA* cluster genes does not induce a symbiotic phenotype. Individual gene deletion strains of the *irlA* cluster genes *pdtA*, *afrA*, *mfsB* and *fxbA* were shown to have normal symbiotic interaction phenotypes in association with *L.*

perenne hosts. This showed that the mechanism responsible for the symbiosis-defective phenotype observed in pDB02-derived Δ*irlA* mutants was not induced by deletion of other genes in this cluster.

The *irlA* cluster genes are upregulated in unfertilised stromata. Through RT-qPCR analysis of material from *D. glomerata* and *Elymus* sp. plants exhibiting *E. typhina* and *E. elymi* stromata, respectively, it was shown expression of all *irlA* cluster genes is significantly upregulated in unfertilised stromata from these associations. This was consistent with RNAseq results provided by our collaborator Prof. Chris Schardl, which showed that expression of the *irlA* cluster genes was also upregulated in unfertilised *E. festucae* stromata. Expression of *irlA* cluster genes in fertilised stromata was either similar to in vegetative tissues or reduced relative to unfertilised stromata depending on the gene tested.

The *irlA* cluster has a proposed role in stromata development. Based on the symbiosis-defective phenotype of Δ*irlA*/KD::*mfsB* strains, stromata-specific expression of all *irlA* cluster genes except *irlA* itself, and similarity to a previously characterised gene cluster shown to be required for sclerotia formation in *Aspergillus flavus* (Forseth et al., 2013), the *irlA* cluster is predicted to have a role in stromata development. Specifically, the *irlA* cluster gene products are proposed to produce a product that promotes hyphal proliferation at the onset of stromata development, while IrlA itself is proposed to have a role in repressing this proliferative signal during vegetative growth.

8 Future work

8.1 Future work based on the results of Chapters 3 & 4

The most obvious future work deriving from the results presented in Chapters 3 and 4 is completion of the experiments designed to identify any novel dipeptide produced by the *E. festucae* E2368 PerA-ΔR protein, as well as analysis of the product release mechanism. At this stage it has not been established whether the *E. festucae* TEII TesB is involved in release of any potential product; however, this will be addressed by the experiments already underway. Radioisotope or stable isotope labelling could also potentially be combined with *perA* and *perA-ΔR*-expressing *P. paxilli* strains generated in this study to conclusively define the substrates of PerA. Such approaches might also aid identification of any novel PerA-ΔR dipeptide product through labelling of the arginine substrate.

Identification of a novel dipeptide produced by PerA-ΔR proteins would be an exciting result, and would represent a nice case study for transposon-driven evolution of novel gene products. In the event that the *P. paxilli* expression approach fails, novel product identification could also be addressed through purification of the PerA-ΔR A1-domain for use in an ATP-pyrophosphate (ATP-PP_i) exchange assay. This assay would be used to identify the amino acid substrate of this domain, and thus infer the final product of PerA-ΔR. ATP-PP_i exchange assays monitor the conversion of ATP to AMP + PP_i, and are commonly used with NRPS A-domains, which catalyse this reaction only when the amino acid substrate is present (Forseth et al., 2013; Lautru and Challis, 2004; Pavela-Vrancic et al., 1994; Schneider et al., 2008). It should be noted that the PerA and PerA-ΔR A1-domains are likely to require the 250 aa N-terminal extending domain to enable function (Kalb et al., 2015).

It would also be interesting to investigate variants such as the PerA-ΔR protein from *E. baconii* to determine if further product diversity from PerA or PerA-ΔR proteins exists. Analysis of the *E. baconii* PerA-ΔR protein could also shed light on the evolution of *perA-ΔR* alleles following the apparently transposon-driven deletion of the sequence encoding the R-domain, as these *E. baconii* proteins do not contain the A1-domain substrate-binding changes found in other PerA-ΔR proteins. Future surveys might also identify a *perA-ΔR* allele in the *E. baconii*, *E. bromicola* or *E. festucae* lineages that has not undergone recombination with a *perA* allele. This

would allow more robust analysis of the trans-species polymorphism hypothesis for the *perA-ΔR* allele. Such an allele would also address the hypothesis that the observed *perA*/*perA-ΔR* recombination events in the *E. baconii*, *E. bromicola* and *E. festucae* lineages was a mechanism to repair a faulty *perA-ΔR* allele.

Based on the observation that a *perA-ΔR* allele is present in an *Epichloë* strain with unusually high—and potentially unexplained—anti-insect activity, deletion of the *perA-ΔR* allele in a suitable strain could enable insect feeding-choice assays to be performed. This would determine whether PerA-ΔR produces an anti-insect compound, and would not be dependent on prior identification of a PerA-ΔR product. If *perA-ΔR*-derived anti-insect activity was detected it would have major implications for the selection of endophyte strains for agricultural purposes. The normal *perA* allele has traditionally been desirable due to the established link to production of the insect-deterring metabolite peramine. If PerA-ΔR were to also produce a product that is bioactive against insect pests, this dogma would have to be revised. Identification of a hybrid strain that expresses functional versions of both PerA and PerA-ΔR would also be particularly desirable. The discovery of anti-insect properties for PerA-ΔR proteins would also have interesting implications for the *perA-ΔR* TSP hypothesis, as maintenance of TSP requires balancing selection. Protection against different groups of insect herbivores to peramine could hypothetically provide such a balancing mechanism.

Finally, it would be interesting to apply the results and observations from these chapters to other NRPSs, and even other SM genes, from within *Epichloë* and across the Clavicipitaceae and filamentous fungi more generally. Specifically, it would be interesting to investigate how commonly recombination occurs between different alleles of NRPS-encoding genes, or even between different NRPS genes entirely. The implications that such recombination events have on protein function and evolution could also be analysed. It would similarly be interesting to analyse the occurrence of TSP in genes responsible for the production of bio-protective metabolites, and investigate if the analogy where these SM genes are compared to an innate immune system is valid.

8.2 Future work based on the results of Chapters 5 & 6

The work presented in these two chapters is entering an interesting phase. A hypothesis for the function of the *irlA* cluster has been established, and resources generated during this study will enable future experiments to investigate the validity of this hypothesis. To further investigate the symbiosis-defective phenotype observed in pDB02-derived tandem-insertion Δ*irlA* strains, an *irlA*/*mfsB* double deletion mutant could be generated. This would test the hypothesis that a Δ*irlA*/KD::*mfsB* genotype is responsible for the observed host stunting and hyphal proliferation *in planta* for pDB02-derived Δ*irlA* strains, as this has not yet been conclusively proven. It may also be possible to complement Δ*irlA*-62 by through a 'knock-in' integration to restore the wt *irlA* locus. If successful, this would prove that the original tandem integration of the *irlA* deletion construct at the *irlA* locus was the cause of the observed symbiotic phenotype. Ectopic expression of *irlA* under control of a constitutively active, highly expressed promoter to help overcome position-specific effects on gene expression might also achieve complementation of the Δ*irlA*-62 phenotype.

The symbiosis-defective Δ*irlA*-62 phenotype is predicted to result from accumulation of an *irlA* cluster product that promotes hyphal proliferation. Confirming that the pDB02-derived Δ*irlA* symbiosis-defective phenotype results from altering the *irlA* locus is therefore important because of the associated implications for *irlA* cluster functionality as a whole. Individual deletions of *pdtA*, *afrA* and *fxbA* within a Δ*irlA*-62 background could also be utilised to gain insight into their function within the *irlA* cluster. If any of these gene deletion mutants complemented the Δ*irlA*-62 symbiosis-defective phenotype it would suggest that production of the proposed proliferation-promoting product had been eliminated. This would support the proposed model for IrlA and *irlA* cluster function where IrlA catalyses conversion of one of the final pathway intermediates into a non-proliferative product.

Comparison of stromata formation frequencies in vernalised plants infected with either wt *E. festucae* Fl1 or the Δ*irlA*-487 strain would test the hypothesis that *irlA* is required to suppress the proliferative signal associated with the onset of stroma formation. It is predicted that Δ*irlA*-487-infected plants would exhibit

significantly higher stromata formation frequency than wt-infected plants. This hypothesis could also be addressed by over-expression of *irlA* in an *Epichloë* strain that frequently forms stromata, as this would be predicted to significantly lower stromata formation frequency. Deletion of the *irlA* cluster genes *pdtA*, *afrA*, *fxbA* and perhaps *mfsB* would also be predicted to significantly lower stromata formation frequency or impair stroma development. This is because *pdtA*, *afrA* and *fxbA* are predicted to encode proteins involved in synthesis of a proliferation-promoting product. Deletion strains for these genes have already been generated in *E. festucae* Fl1. However, Fl1 only forms stromata infrequently. This means that a large sample size would be required to identify any significant reduction in stromata formation frequency in plants infected with Δ*pdtA*, Δ*afrA*, Δ*mfsB* or Δ*fxbA* strains relative to wt-infected plants. It would instead be advisable to generate *irlA* cluster gene deletions in *Epichloë* strains that readily form stromata, such as many *E. typhina* and *E. elymi* isolates. Associations involving such gene deletion strains could then be utilised to more efficiently assay for reduction of stromata formation frequency or impairment of stroma development.

In the event that the *irlA* cluster is confirmed to be involved in initiation of hyphal proliferation prior to stroma development, it would be interesting to investigate how the regulation of this cluster controls entry into this pathway. RT-qPCR or RNAseq analyses could be used to determine the temporal and spatial profiles of *irlA* cluster gene expression during stromata development. This would determine when and where *irlA* cluster gene expression is required during stroma development. Expression of *irlA* cluster genes could also be analysed in vegetative tissues from different associations that exhibit diverse different stromata formation frequencies. This data could be used to address the hypothesis that differences in basal expression of *irlA* cluster genes might control the frequency with which an isolate forms stromata.

Finally, research efforts should be focused on identification of product(s) produced by the proteins encoded by the *irlA* cluster. This would enable targeted metabolite analyses for the *irlA* cluster products in material from plant-endophyte associations to better understand the roles that these products play in the symbiosis.

Product identification would also provide considerable insight into the *irlA* cluster biosynthetic pathway as a whole, and pathways involving AFRs more generally.

An important initial step in product identification would be determination of the amino acid substrate modified by AfrA, as AfrA is predicted to catalyse the first biosynthetic step in this pathway. This substrate amino acid would therefore be predicted to provide the basic structure upon which the final metabolite product(s) would be synthesised. Identification of the AfrA substrate could be addressed using the same method employed in previous studies of fungal AFRs (Forseth et al., 2013; Schneider et al., 2008). The whole AfrA protein would be expressed in a suitable heterologous host (e.g. *E. coli* or *S. cerevisiae*), then purified and used in an ATP-PP_i exchange assay to identify the amino acid substrate. This approach avoids potential protein-folding problems that can be encountered when A-domains are expressed outside of their native context within a larger NRPS/AFR (Kalb et al., 2015; Lee, 2008). Expression of the entire AfrA protein does not complicate the ATP-PP_i assays as, unlike most regular NRPSs, AFR proteins only contain a single A-domain and therefore usually only bind a single amino acid substrate. Once the amino acid substrate is identified, purified AfrA protein could also be analysed *in vitro* to determine the final product. AfrA would require modification by a fungal 4'PPTase *in vitro* to become functional, but a suitable system for this process has already been established in a previous study (Schneider et al., 2008). Knowledge of this base metabolite in the *irlA* cluster biosynthetic pathway might also facilitate identification of derivative metabolites from stromata. The functionality of other *irlA* cluster proteins could potentially also be analysed *in vitro*.

An alternative or complementary approach could be to utilise the heterologous *P. paxilli* expression system established in this study for expression of the *irlA* cluster genes. This would involve expression of *irlA* cluster genes under the control of different *pax* promoters to identify resulting products, as well as expression of gene subsets from the *irlA* cluster to characterise the function of individual proteins within the overall biosynthetic pathway. Identification of the AfrA amino acid substrate would aid this approach, as it would provide significant insight into the likely structure and chemistry of the products being searched for, but is not an absolute requirement. However, if the AfrA substrate was known, approaches

such as radioisotope or stable isotope labelling of this amino acid could be utilised to aid product identification.

9 Appendices

9.1 Accompanying digital media

- Raw peramine concentration data
- Multiple sequence alignment of *tesB* across *Epichloë*.
- Multiple sequence alignment of *irlA* across *Epichloë*.
- Multiple sequence alignment of *pdtA* across *Epichloë*.
- Multiple sequence alignment of *afrA* across *Epichloë*.
- Multiple sequence alignment of *mfsB* across *Epichloë*.
- Multiple sequence alignment of *fxbA* across *Epichloë*.
- Raw C_p values from RT-qPCR analysis of RNA extracted from *E. typhina*-infected *D. glomerata* plants.
- Raw C_p values from RT-qPCR analysis of RNA extracted from *E. elymi*-infected *Elymus* sp. plants.
- Raw C_p values from qPCR analysis of DNA extracted from *E. typhina*-infected *D. glomerata* plants.
- Raw C_p values from RT-qPCR analysis of RNA extracted from *L. perenne* plants infected with *E. festucae* wt, Δ*irlA*-62, Δ*irlA*-133, Δ*irlA*-438, Δ*irlA*-487 or Δ*irlA*-573.
- Raw C_p values from RT-qPCR analysis of RNA extracted from mycelia of *E. festucae* wt, Δ*irlA*-62, Δ*irlA*-133, Δ*irlA*-438, Δ*irlA*-487 or Δ*irlA*-573 grown in culture.

9.2 Supplementary material for Chapter 3

Table S1. Nonhybrid *Epichloë* species used in this study

Endophyte species	Isolate ID	ATCC or CBS number	Host species	Reference or source
E. amarillans	E52	ATCC 200743	*Sphenopholis obtusata*	(1-4)
E. amarillans	E57	ATCC 200744	*Agrostis hyemalis*	(1-5)
E. amarillans	E4668			(6)
E. baconii	As6	ATCC 90167	*Agrostis capillaris*	(7)
E. baconii	E424	ATCC 200746	*Agrostis tenuis*	(8)
E. baconii	E357 = E1031	ATCC 200745	*Calamagrostis villosa*	(2-4, 6, 8)
E. brachyelytri	E1040	ATCC 200752	*Brachyelytrum erectum*	(2-4)
E. brachyelytri	E1045	ATCC 200753	*Brachyelytrum erectum*	(3)
E. brachyelytri	E1124	ATCC 201560	*Brachyelytrum erectum*	(3)
E. brachyelytri	E4804		*Brachyelytrum erectum*	(5)
E. bromicola	E501	ATCC 200749	*Bromus erectus*	(2, 4, 8)
E. bromicola	E502	ATCC 200750	*Bromus erectus*	(5, 8)
E. bromicola	E799	ATCC 201559	*Bromus benekenii*	(8)
E. elymi	E56	ATCC 201551	*Elymus canadensis*	(1-5, 9)
E. elymi	E184	ATCC 200850	*Elymus virginicus*	(2-4, 9)
E. elymi	EC1		*Elymus canadensis*	S. Marek, Oklahoma State University
E. elymi	EC4		*Elymus canadensis*	S. Marek, Oklahoma State University
E. elymi	WWG2		*Pascopyrum smithii*	C.A. Young, The Samuel Roberts Noble Foundation, Oklahoma
E. festucae	E189	ATCC 90661	*Festuca rubra*	(1, 4, 8, 9)
E. festucae	E434	ATCC MYA-434	*Festuca gigantea*	(2, 8)
E. festucae	E1035.33	ATCC MYA-446	NA	(10)
E. festucae	E2368		NA	(5, 10)
E. festucae	Fg1		*Festuca glauca*	(9)
E. festucae	Fl1	ATCC MYA-3407	*Festuca longifolia*	(9, 11)

E. festucae	Frc5		*Festuca rubra* subsp. *commutata*	(9)
E. festucae	Frc7		*Festuca rubra* subsp. *commutata*	(9)
E. festucae	Fr1		*Festuca rubra*	(9, 11)
E. festucae	Frr1		*Festuca rubra* subsp. *rubra*	(9)
E. glyceriae	E2772	ATCC 200755	*Glyceria striata*	(2-4)
E. mollis	E3601 = AL9924		*Holcus mollis*	(6)
E. sylvatica	E354	ATCC 200748	*Brachypodium sylvaticum*	(2-4, 8)
E. sylvatica	E503	ATCC 200751	*Brachypodium sylvaticum*	(2-4, 8)
E. typhina	E8	ATCC 200736	*Lolium perenne*	(3-5, 8, 9)
E. typhina	E505	ATCC 200739	*Brachypodium pinnatum*	(2-4, 8)
E. typhina	ORE04		*Dactylis glomerata*	C.A. Young, The Samuel Roberts Noble Foundation, Oklahoma
E. typhina	ORE06		*Dactylis glomerata*	C.A. Young, The Samuel Roberts Noble Foundation, Oklahoma
E. typhina subsp. *clarkii*	Holcus 3	ATCC 90168	*Holcus lanatus*	(7)
E. typhina subsp. *poae*	E1022	ATCC 201668	*Poa nemoralis*	(4, 8)
E. typhina subsp. *poae*	E5819		*Poa nemoralis*	(5)

Table S2. *Epichloë* species from endophyte-infected plant material used in this study

Species	Isolate ID	Plant ID	Host Species	Reference or source
E. typhina	OR10	Collected from field plant in Oregon	*Dactylis glomerata*	C.A. Young, The Samuel Roberts Noble Foundation, Oklahoma
E. typhina subsp. *poae* BlaTG-1	NFe671	BRLA-671	*Bromus laevipes*	(12)
E. amarillans	NFe-708	ELCA-708	*Elymus canadensis*	C.A. Young, The Samuel Roberts Noble Foundation, Oklahoma
E. elymi	NFE728	ELVI-728	*Elymus virginicus*	C.A. Young, The Samuel Roberts Noble Foundation, Oklahoma
E. elymi	NFE741	ELVI-741	*Elymus virginicus*	C.A. Young, The Samuel Roberts Noble Foundation, Oklahoma
E. festucae var. *lolii*	NFe243	LOPE-243	*Lolium perenne*	C.A. Young, The Samuel Roberts Noble Foundation, Oklahoma
E. cabralii BlaTG-2	NFe661	BRLA-661	*Bromus laevipes*	(12)
E. cabralii BlaTG-2	NFe688	BRLA-688	*Bromus laevipes*	(12)
E. sp. BlaTG-3 (G1)	NFe79	BRLA-79	*Bromus laevipes*	(12)
E. sp. BlaTG-3 (G2)	NFe82	BRLA-82	*Bromus laevipes*	(12)
E. sp. BlaTG-3 (G3)	NFe83	BRLA-83	*Bromus laevipes*	(12)
E. canadensis	NFe699	ELCA-699	*Elymus canadensis*	C.A. Young, The Samuel Roberts Noble Foundation, Oklahoma
E. amarillans	NFe715	ELCA-715	*Elymus canadensis*	C.A. Young, The Samuel Roberts Noble Foundation, Oklahoma
E. canadensis	NFe1001 = NFe727	ELCA-727	*Elymus canadensis*	(13)
E. canadensis	NFe716	ELCA-716	*Elymus canadensis*	(13)
E. coenophiala	e19 = ATCC 90664	LOAR-00437	*Lolium arundinaceum*	(4, 14)

E. coenophiala	AR584	LOAR-00405	*Lolium arundinaceum*	(15, 16)
E. coenophiala	E34	LOAR-00245 (BarOptima PLUS E34, Barenbrug	*Lolium arundinaceum*	C.A. Young, The Samuel Roberts Noble Foundation, Oklahoma
E. coenophiala	NFe45078	LOAR-00190	*Lolium arundinaceum*	(17)
E. sp. FaTG-2 G2	NFe45079	LOAR-00193	*Lolium arundinaceum*	(17)
E. sp. FaTG-2 G3	NFe45115	LOAR-00210	*Lolium arundinaceum*	(17)
E. sp. FaTG-3	NFe1100	LOAR-00488	*Lolium arundinaceum*	(6, 16)
E. sp. FaTG-4	e4305	4305	*Lolium sp.*	(6, 16)
E. uncinata	e167 = CBS 102646	FEPR-00400	*Lolium pratense*	(6, 18-20)
E. siegelii	e915	955	*Lolium pratense*	(20, 21)
E. sp.	e4678	4678	*Poa alsodes*	C.L. Schardl, University of Kentucky
E. sp.	e4742	4742	*Poa autumnalis*	C.L. Schardl, University of Kentucky
E. sp.	e4755	4755	*Poa autumnalis*	C.L. Schardl, University of Kentucky
E. sp. PauTG-1	e55		*Poa autumnalis*	(4)
E. sp.	e4427	4427	*Sphenopholis sp.*	C.L. Schardl, University of Kentucky
E. sp.	e4768	4768	*Festuca versuta*	C.L. Schardl, University of Kentucky
E. sp.	e4672	4672	*Agrostis hyemalis*	C.L. Schardl, University of Kentucky
E. sp.	e4676	4676	*Agrostis hyemalis*	C.L. Schardl, University of Kentucky

Table S3. PCR primers used in this study

Target[1]	Primers used	Product size (bp)[2]
mfsA	mfsA_F: ACGATGGATTGGTCTTCTC	1362
	mfsA_R: GCAGACGAGCAACGCAACG	
qcrA	qcrA_F: TTCCGAGATTGTCGAGG	639
	qcrA_R: TCACTCTCCTGCTGCTG	
perA A1-domain	perA-5'_F3: ATGACGAGCTCGGAGCGAGTTG	1691
	perA-A1_R: AGACTTCCATCTGCACAGTATC	
perA T1-domain	perA1_4: TCGGAAAGGTCGGCTGTAC	1073
	perA1_R: TTGCTTCATCCCAGTCAGC	
perA C-domain	perA2_F: ATCCAAGACGCATATCCC	878
	perA-C_R: ATCATCTCGGCGGCTTCC	
perA A2-domain	perA2_1: ACAGCTTTGCCACTCCAAG	2363
	perA2_R: ATCCACGCCTATGTAGCTC	
perA M-domain	perA3_F: GCTTGCTGCGTTTGTCAC	1298
	perA-M_R: TGGGAAATCGGAACAAGG	
perA T2-domain	perA-T2_F: TCTTCAGGCATCGCAGGAAC	600
	perA-T2_R: TCGGCCACCTCCAGCCTGATG	
perA R*-domain	perA3_3: AGGAAGGCATCAGGCTGG	1376
	perA3_R: CTAGCCTCCAGATCTTGTG	
perA R*-domain deletion	perA-T2_F: TCTTCAGGCATCGCAGGAAC	742 (E2368)
	perA-17bp_R: GTACGGATAACCTCAAC	
perA-1	perA1_F: ATGGACGCGGAGCCTTTTG	2905
	perA1_R: TTGCTTCATCCCAGTCAGC	
	Additional sequencing primers:	
	perA1_1: TTTGCAGTCCGGCGAAGC	
	perA1_2: ACGCCTAGAGTTCTGCAT	
	perA1_3: ATACTCACGCACACCAAC	
	perA1_4: TCGGAAAGGTCGGCTGTAC	

perA-2	perA2_F: ATCCAAGACGCATATCCC	2973
	perA2_R: ATCCACGCCTATGTAGCTC	
	Additional sequencing primers:	
	perA2_1: ACAGCTTTGCCACTCCAAG	
	perA2_2: AGATTTGGCAGTGGAACG	
	perA2_3: GTCTCCATCAGCAACATC	
perA-3	perA3_F: GCTTGCTGCGTTTGTCAC	3058
	perA3_R: CTAGCCTCCAGATCTTGTG	
	Additional sequencing primers:	
	perA3_1: TATCTCTTCAACATCATCCAG	
	perA3_2: ATCGCAGGAACACTCATCG	
	perA3_3: AGGAAGGCATCAGGCTGG	
	perA3_4: ATTCGCCAGGATGTAGAG	
perA-3ΔR*	perA3_F: GCTTGCTGCGTTTGTCAC	1864
	perA-17bp_R: GTACGGATAACCTCAAC	(E2368)
	Additional sequencing primers:	
	perA3_1: TATCTCTTCAACATCATCCAG	
	perA3_2: ATCGCAGGAACACTCATCG	
	perA3_3: AGGAAGGCATCAGGCTGG	
perA-qcrA	perA-mid2_F2: CATCAGGCTGGAGGTGGCCGA	1451
intergenic region	Ef104_R: CTAAGCTTTGGTCCAAGCTGCG	(E2368)

[1] Abbreviations: A = adenylation, T = thiolation, C = condensation, M = methylation and R* = reductase*.

[2] Product fragment sizes are for amplification of *E. festucae* Fl1 genomic DNA unless *E. festucae* E2368 is indicated in brackets.

Table S4. Details of sequenced *perA* alleles

Species	Isolate	Allele progenitor[a]	*perA* type	Notes[b]	Accession #[c]
Nonhybrid species					
E. amarillans	E57	*E. amarillans*	*perA*		**JN640285**
E. amarillans	E4668	*E. amarillans*	*perA*		
E. baconii	As6	*E. baconii*	*perA*-ΔR		KP347847
	E357/E1031	*E. baconii*	*perA*		KP347874
	E424	*E. baconii*	*perA*-ΔR		KP347851
E. brachyelytri	E4804	*E. brachyelytri*	*perA*		**JN613323**
E. bromicola	AL0434	*E. bromicola*	*perA*		KP347845
	AL0426_2	*E. bromicola*	*perA*-ΔR	START	KP347846, KP719965
	E501	*E. bromicola*	*perA*		KP347854
	E502	*E. bromicola*	*perA*		**JX441995**
	E799	*E. bromicola*	*perA*-ΔR	STOP@7051	KP347855
E. elymi	E56	*E. elymi*	*perA*		**JX402755**
E. festucae	E189	*E. festucae*	*perA*-ΔR		KP347868
	E2368	*E. festucae*	*perA*-ΔR		**JN640287**
	Fg1	*E. festucae*	*perA*	-236@3652	KP347869
	Fl1	*E. festucae*	*perA*		**AB205145**
E. mollis	E3601	*E. mollis*	*perA*	+1@1412, +2@3470, +4@6153, +7@6162	KP347873
E. sylvatica	E354	*E. sylvatica*	*perA*-ΔR	+1@992, -259@1891, +2@4200	KP347850, KP719966
	E503	*E. sylvatica*	*perA*-ΔR	+1@992, -259@1891, +2@4200	KP347852
E. typhina	E8	*E. typhina*	*perA*		**JX402754**
	E505	*E. typhina*	*perA*-ΔR	STOP@2116	KP347853, KP719967
	E1022	*E. typhina*	*perA*-ΔR		KP347856
	ORE04	*E. typhina*	*perA*-ΔR	STOP@7045	KP347848, KP719969
	ORE06	*E. typhina*	*perA*-ΔR	STOP@7045	KP347849
	OR10	*E. typhina*	*perA*-ΔR	incomplete	KP347870
E. typhina subsp. *clarkii*	Holcus 3	*E. typhina* subsp. *clarkii*	*perA*-ΔR	STOP@6628	KP347859, KP719968
E. typhina subsp. *poae*	BlaTG-1	*E. typhina* subsp. *poae*	*perA*	incomplete	KP347875
	E5819	*E. typhina* subsp. *poae*	*perA*-ΔR		**JN640289**, KP719970

Hybrid species					
E. cabralii	BlaTG-2 BRLA-00661	*E. amarillans*	*perA*	incomplete	KP347876
		E. typhina	*perA*	incomplete	KP347877
E. siegelii	e915	*E. festucae*	*perA*-ΔR		KP347857, KP719971-KP719973
		E. bromicola	*perA*-ΔR	STOP@2839	KP347858
E. uncinata	e167	*E. bromicola*	*perA*	-1@5044	KP347860
		E. typhina	*perA*	+1@7972	KP347861
E. sp. FaTG-2 G2	NFe45079	LAE	*perA*	+1@1031, -621@1251, -328@3970	KP347862
		E. festucae	*perA*		KP347863
E. sp. FaTG-2 G3	NFe45115	LAE	*perA*	+1@1031, -621@1251, -328@3970	KP347864
		E. festcuae	*perA*	STOP@525	KP347865
E. sp. FaTG-3	NFe1100	LAE	*perA*	+1@1031, -621@1251, -328@3970	KP347866
		E. typhina	*perA*		KP347867
E. sp.	e4768	*E. festucae*	*perA*-ΔR	incomplete	KP347871
		E. typhina	*perA*	incomplete	KP347872

[a]LAE = *Lolium*-associated endophyte

[b]The sign indicates if bases have been deleted (-) or inserted (+), the following number shows the number of bp affected and the number after the "@" symbol shows the coordinates at which, or immediately before which, the insertion or deletion begins, respectively. Small deletions or insertions that do not cause a frameshift mutation are not detailed here. The sizes of large deletions are given relative to the *perA* sequence from *E. festucae* Fl1. STOP indicates a nonsense mutation, with the number following "@" detailing the coordinate of the causative SNP. START indicates the ATG start codon has been disrupted, though a nearby ATG codon may fill this role.

[c]Accession numbers shown in bold are from previous publications (5, 22).

References for supplementary material only

1. **Schardl CL, Leuchtmann A, Tsai HF, 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.

2. **Schardl CL, Leuchtmann A, Chung KR, Penny D, Siegel MR.** 1997. Coevolution by common descent of fungal symbionts (Epichloe spp) and grass hosts. Mol. Biol. Evol. **14:**133-143.

3. **Schardl CL, Leuchtmann A.** 1999. Three new species of Epichloe symbiotic with North American grasses. Mycologia **91:**95-107.

4. **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.

5. **Schardl CL, Young CA, Hesse U, Amyotte SG, Andreeva K, Calie PJ, Fleetwood DJ, Haws DC, Moore N, Oeser B, Panaccione DG, Schweri KK, Voisey CR, Farman ML, Jaromczyk JW, Roe BA, O'Sullivan DM, Scott B, Tudzynski P, An Z, Arnaoudova EG, Bullock CT, Charlton ND, Chen L, Cox M, Dinkins RD, Florea S, Glenn AE, Gordon A, Gueldener U, Harris DR, Hollin W, Jaromczyk J, Johnson RD, Khan AK, Leistner E, Leuchtmann A, Li C, Liu J, Liu J, Liu M, Mace W, Machado C, Nagabhyru P, Pan J, Schmid J, Sugawara K, Steiner U, Takach JE, Tanaka E, Webb JS, Wilson EV, Wiseman JL, Yoshida R, Zeng Z.** 2013. Plant-symbiotic fungi as chemical engineers: Multi-genome analysis of the Clavicipitaceae reveals dynamics of alkaloid loci. PLoS Genet. **9**.

6. **Schardl CL, Young CA, Pan J, Florea S, Takach JE, Panaccione DG, Farman ML, Webb JS, Jaromczyk J, Charlton ND, Nagabhyru P, Chen L, Shi C, Leuchtmann A.** 2013. Currencies of mutualisms: Sources of alkaloid genes in vertically transmitted epichloae. Toxins **5:**1064-1088.

7. **White JF.** 1993. Endophyte-host associations in grasses .19. A systematic study of some sympatric species of epichloe in england. Mycologia **85:**444-455.

8. **Leuchtmann A, Schardl CL.** 1998. Mating compatibility and phylogenetic relationships among two new species of Epichloe and other congeneric European species. Mycol. Res. **102:**1169-1182.

9. **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.

10. **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.

11. **Leuchtmann A.** 1994. Isozyme relationships of *Acremonium* endophytes from 12 *Festuca* species. Mycol. Res. **98:**25-33.

12. **Charlton ND, Craven KD, Afkhami ME, Hall BA, Ghimire SR, Young CA.** 2014. Interspecific hybridization and bioactive alkaloid variation increases diversity in endophytic *Epichloë* species of *Bromus laevipes*. FEMS Microbiol. Ecol.**:**doi: 10.1111/1574-6941.12393.

13. **Charlton ND, Craven KD, Mittal S, Hopkins AA, Young CA.** 2012. *Epichloë canadensis*, a new interspecific epichloid hybrid symbiotic with Canada wildrye (*Elymus canadensis*). Mycologia **104:**1187-1199.

14. **Tsai HF, Siegel MR, Schardl CL.** 1992. Transformation of *Acremonium coenophialum*, a protective fungal symbiont of the grass *Festuca arundinacea*. Current Genetics **22:**399-406.

15. **Johnson LJ, de Bonth ACM, Briggs LR, Caradus JR, Finch SC, Fleetwood DJ, Fletcher LR, Hume DE, Johnson RD, Popay AJ, Tapper BA, Simpson WR, Voisey CR, Card SD.** 2013. The exploitation of epichloae endophytes for agricultural benefit. Fungal Diversity **60:**171-188.

16. **Takach JE, Young CA.** 2014. Alkaloid genotype diversity of tall fescue endophytes. Crop Sci. **54:**667-678.

17. **Takach JE, Mittal S, Swoboda GA, Bright SK, Trammell MA, Hopkins AA, Young CA.** 2012. Genotypic and chemotypic diversity of *Neotyphodium* endophytes in tall fescue from Greece. Appl. Environ. Microbiol. **78:**5501-5510.

18. **Gams W, Petrini O, Schmidt D.** 1990. *Acremonium uncinatum*, a new endophyte in *Festuca pratensis*. Mycotaxon **37:**67-71.

19. **Glenn AE, Bacon CW, Price R, Hanlin RT.** 1996. Molecular phylogeny of Acremonium and its taxonomic implications. Mycologia **88:**369-383.

20. **Leuchtmann A, Bacon CW, Schardl CL, White JF, Jr., Tadych M.** 2014. Nomenclatural realignment of *Neotyphodium* species with genus *Epichloë*. Mycologia **106:**202-215.

21. **Craven KD, Blankenship JD, Leuchtmann A, Hignight K, Schardl CL.** 2001. Hybrid fungal endophytes symbiotic with the grass Lolium pratense. Sydowia **53:**44-73.

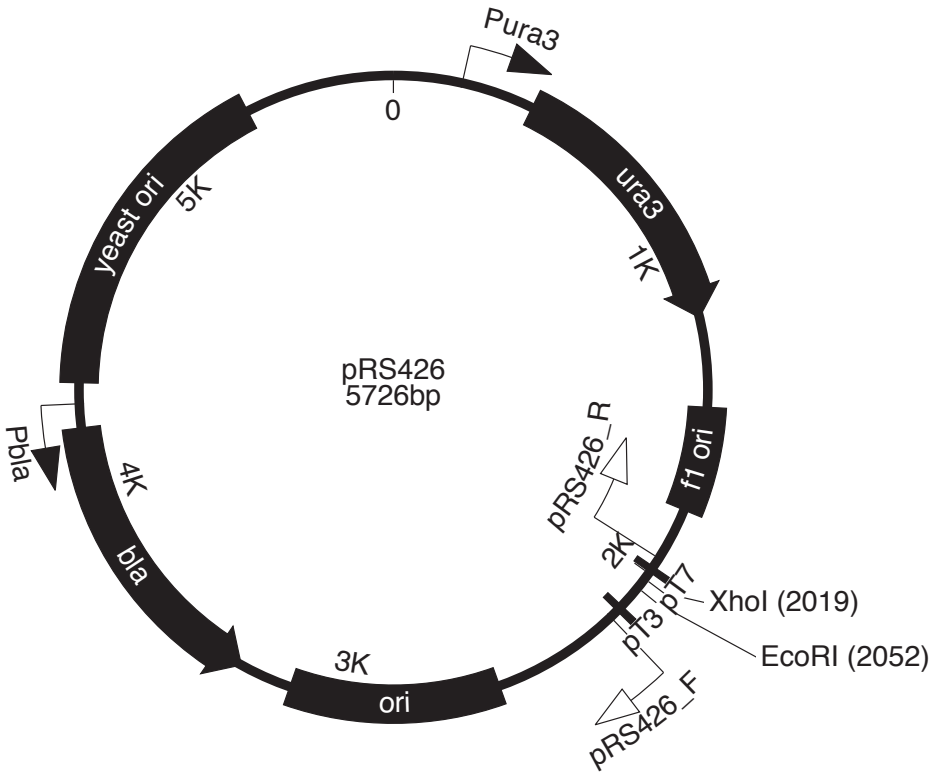
22. **Tanaka A, Tapper BA, Popay A, Parker EJ, Scott B.** 2005. A symbiosis expressed non-ribosomal peptide synthetase from a mutualistic fungal endophyte of perennial ryegrass confers protection to the symbiotum from insect herbivory. Mol. Microbiol. **57:**1036-1050.

9.3 Plasmid gene maps

Note: All PCR reactions generating DNA fragments for plasmid assembly were performed using high-fidelity enzymes (2.6.4). Plasmid inserts were sequenced (2.6.13) to confirm the sequence was free of potentially deleterious PCR-induced mutations.

9.3.1 Details of plasmid pRS426

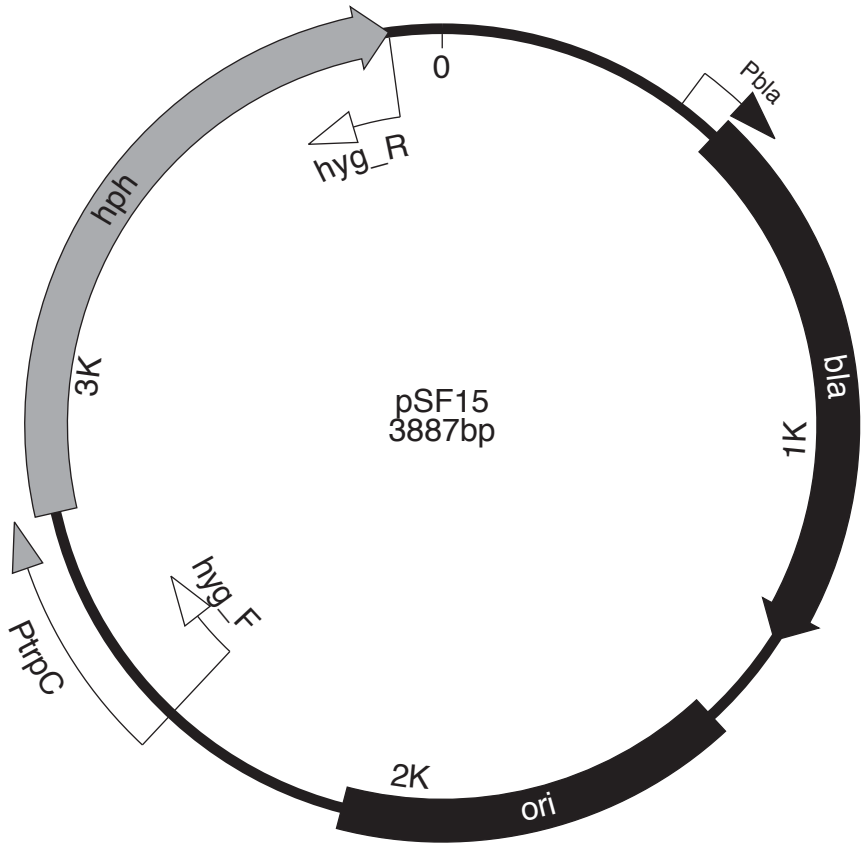
Purpose: General purpose cloning vector for replication in *E. coli*, *S. cerevisiae* and f1 phage. Specifically designed for use with yeast recombinational cloning (2.6.11).
Plasmid map:



9.3.2 Details of plasmid pSF15.15

Purpose: Co-transformation partner plasmid to provide hygromycin resistance for transformant screening. Also used as template for PCR amplification of P*trpC–hph* fragment for assembly of several pDB plasmids.

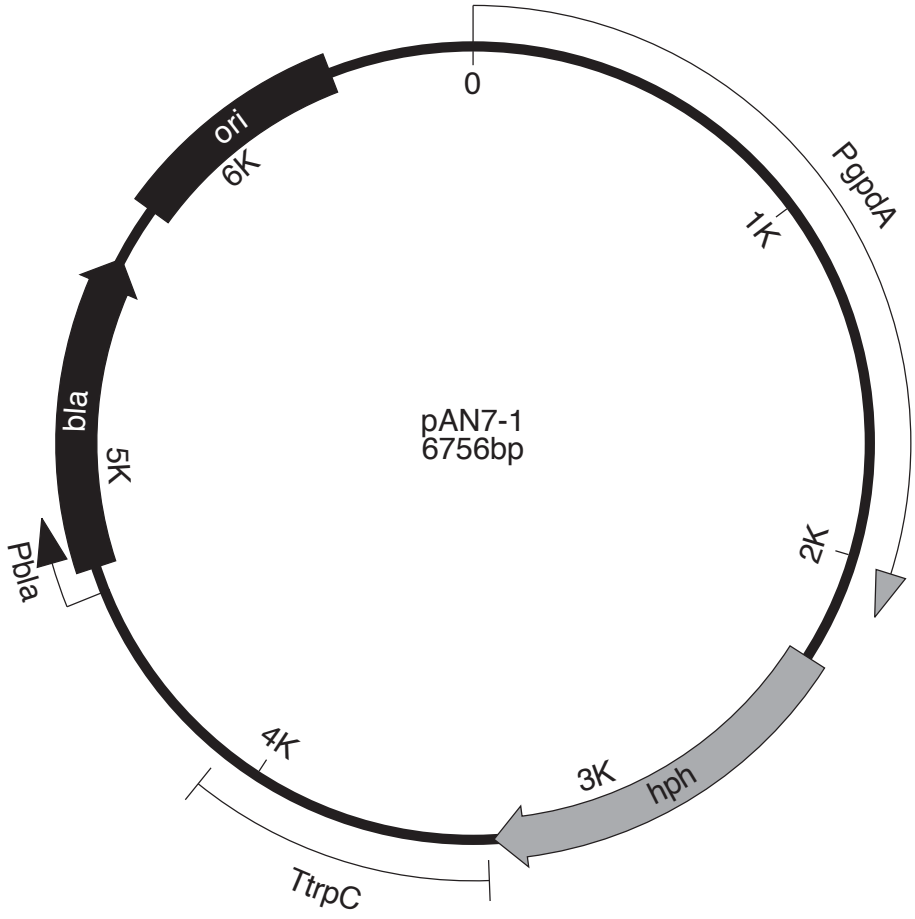
Plasmid map:



9.3.3 Details of plasmid pAN7-1

Purpose: Co-transformation partner plasmid to provide hygromycin resistance for transformant screening.

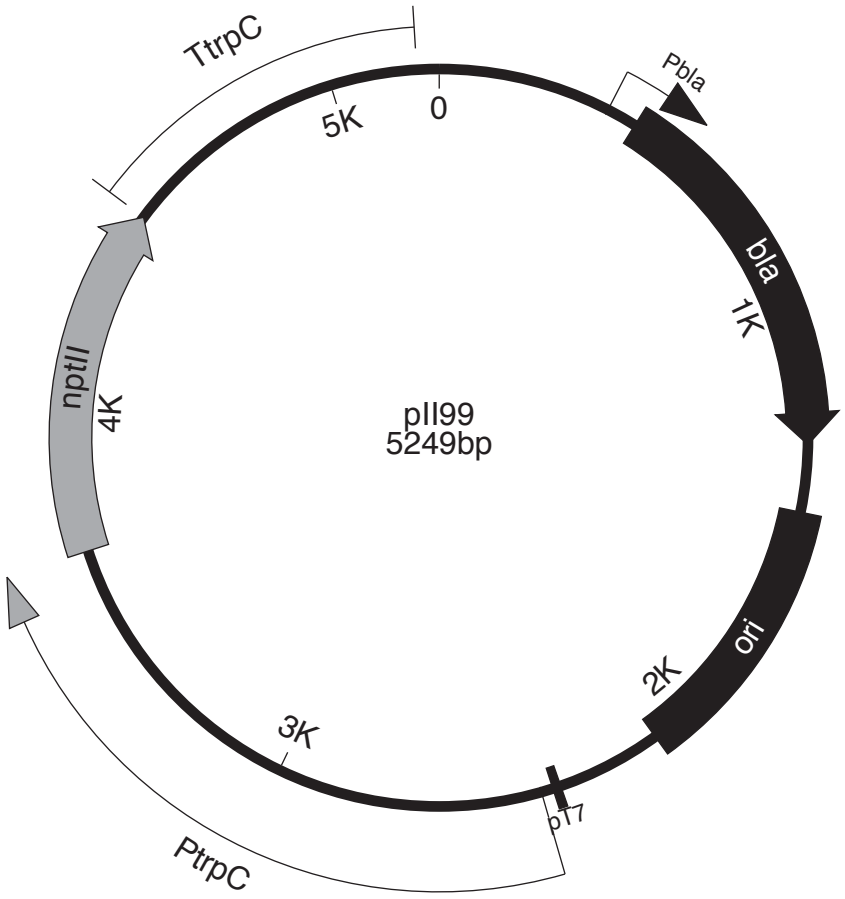
Plasmid map:



9.3.4 Details of plasmid pII99

Purpose: Co-transformation partner plasmid to provide hygromycin resistance for transformant screening.

Plasmid map:

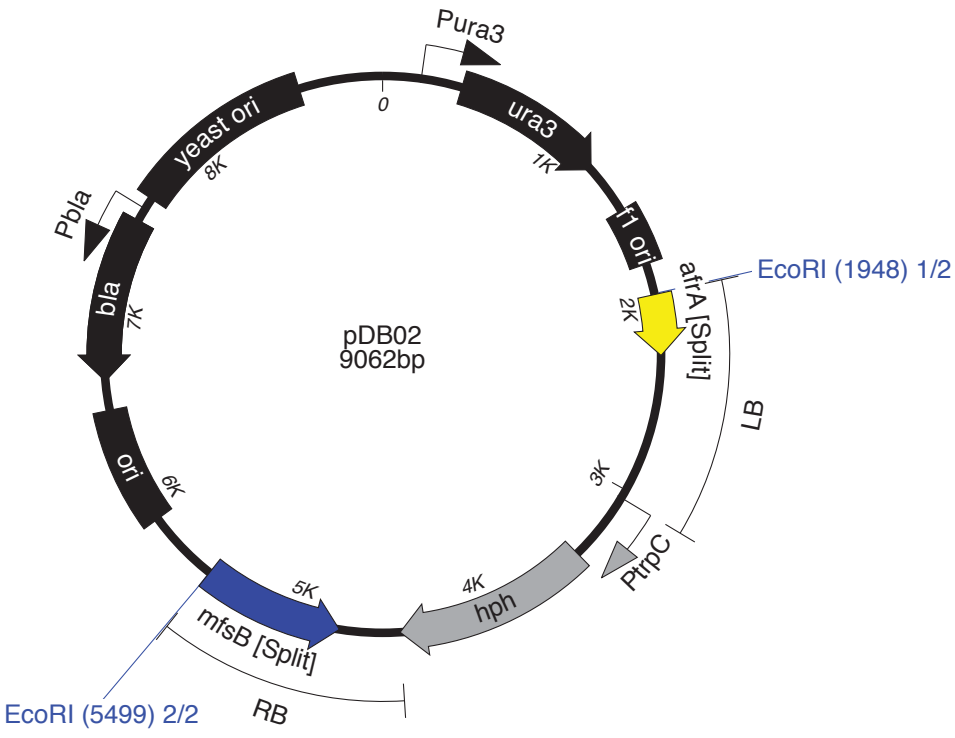


9.3.5 Details of plasmid pDB02

Purpose: *irlA* deletion construct
Insert: *irlA* LB–P*trpC-hph–irlA* RB
Vector: pRS426
Assembly fragments:

1. 5693 bp pRS426 backbone fragment generated by digestion of pRS426 with *Xho*I and *Eco*RI.
2. 1154 bp *irlA* upstream fragment PCR amplified from *E. festucae* Fl1 gDNA template using primer pair pDB02_1 / pDB02_2.
3. 1394 bp P*trpC-hph* fragment PCR amplified from pSF15.15 template using primer pair hyg_F / hyg_R.
4. 1122 bp *irlA* downstream fragment PCR amplified from *E. festucae* Fl1 gDNA template using primer pair pDB02_3 / pDB02_4.

Assembly method: Yeast recombinational cloning (2.6.11)
Sequencing: Fidelity of insert sequence confirmed 100%
Plasmid map:



9.3.6 Details of plasmid pDB05

Purpose: Heterologous expression of *perA* in *P. paxilli*

Insert: P*paxM*–*perA*–T*paxM*

Vector: pRS426

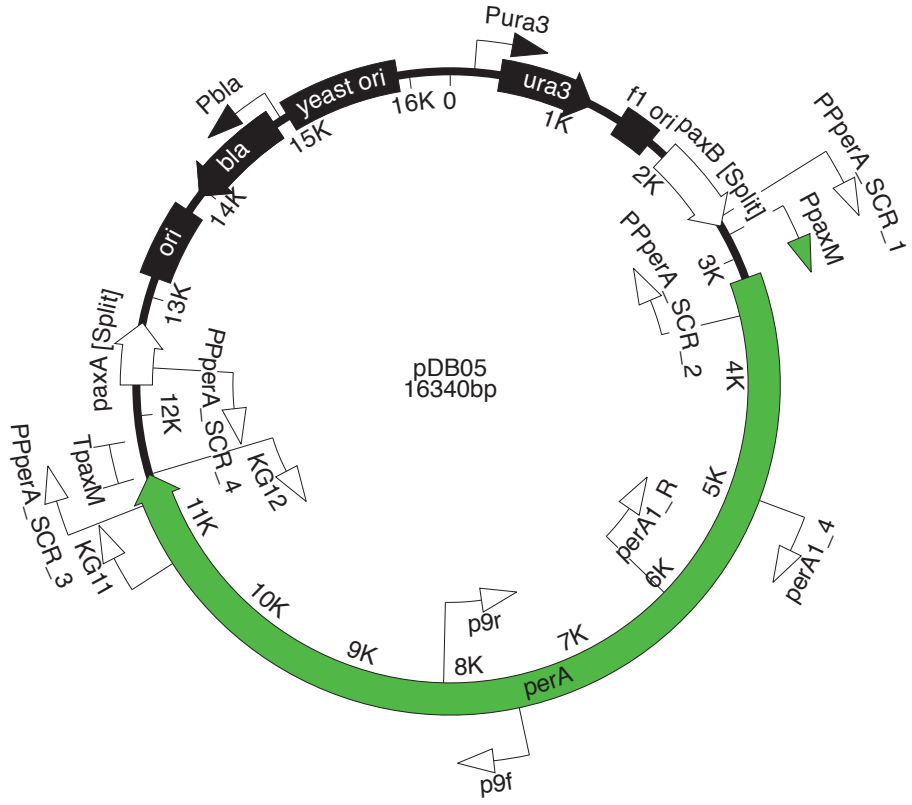
Assembly fragments:

1. 5693 bp pRS426 backbone fragment generated by digestion of pRS426 with *Xho*I and *Eco*RI.
2. 1292 bp P*paxM* fragment PCR amplified from *P. paxilli* gDNA using primer pair pDB05_1/pDB05_2.
3. 2459 bp *perA* fragment PCR amplified from *E. festucae* Fl1 gDNA using primer pair pDB05_3/pDB05_4.
4. 3085 bp *perA* fragment PCR amplified from *E. festucae* Fl1 gDNA using primer pair pDB05_5/pDB05_6.
5. 3079 bp *perA* fragment PCR amplified from *E. festucae* Fl1 gDNA using primer pair pDB05_7/pDB05_8.
6. 1336 bp P*paxM* fragment PCR amplified from *P. paxilli* gDNA using primer pair pDB05_9/pDB05_10.

Assembly method: Yeast recombinational cloning (2.6.11)

Sequencing: Fidelity of insert sequence confirmed 100%

Plasmid map:



9.3.7 Details of plasmid pDB08

Purpose: Complementation of Δ*irlA* mutants.

Insert: P*irlA*–*irlA*–T*irlA*

Vector: pRS426

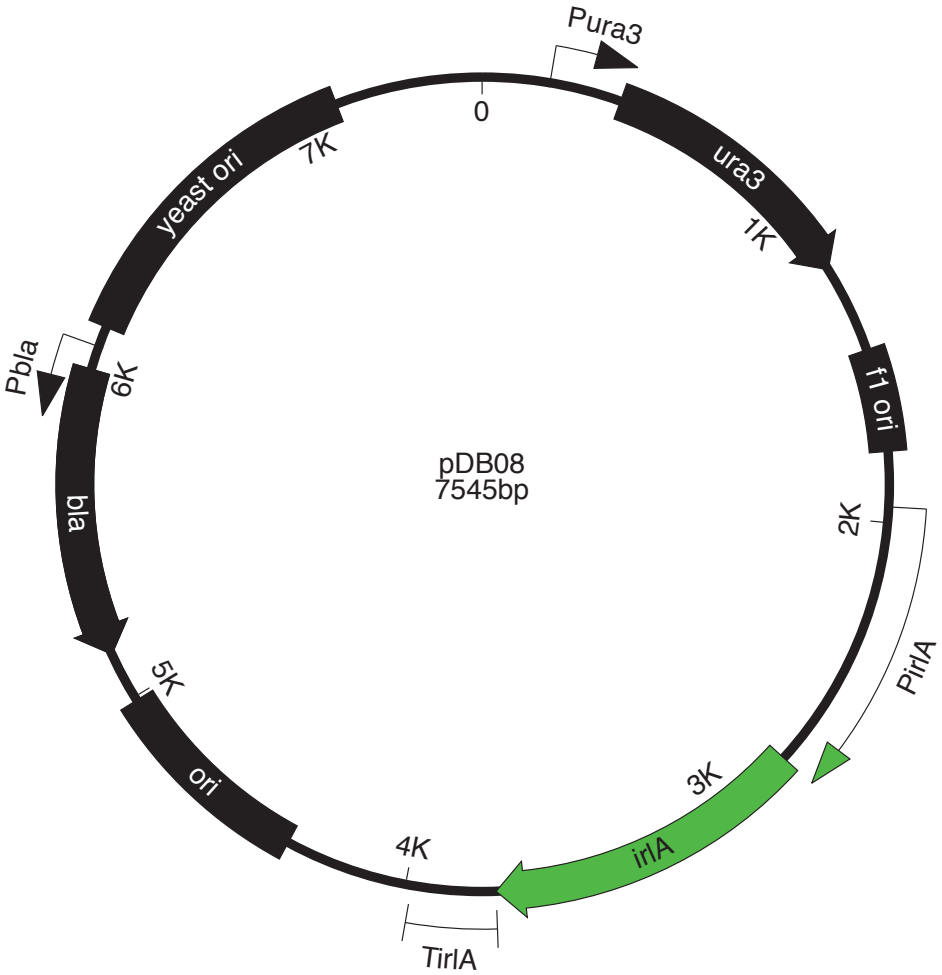
Assembly fragments:

1. 5505 bp vector backbone fragment amplified from linearised pRS426 using primer pair pRS426_F/pRS426_R.
2. 2097 bp P*irlA*–*irlA*–T*irlA* fragment amplified from *E. festucae* Fl1 gDNA using primer pair pDB08_1/pDB08_2.

Assembly method: Gibson assembly (2.6.12)

Sequencing: Fidelity of insert sequence confirmed 100%

Plasmid map:



9.3.8 Details of plasmid pDB09

Purpose: Investigate localisation of *irlA*

Insert: P*irlA*–*irlA*-*eGFP*–T*irlA*

Vector: pRS426

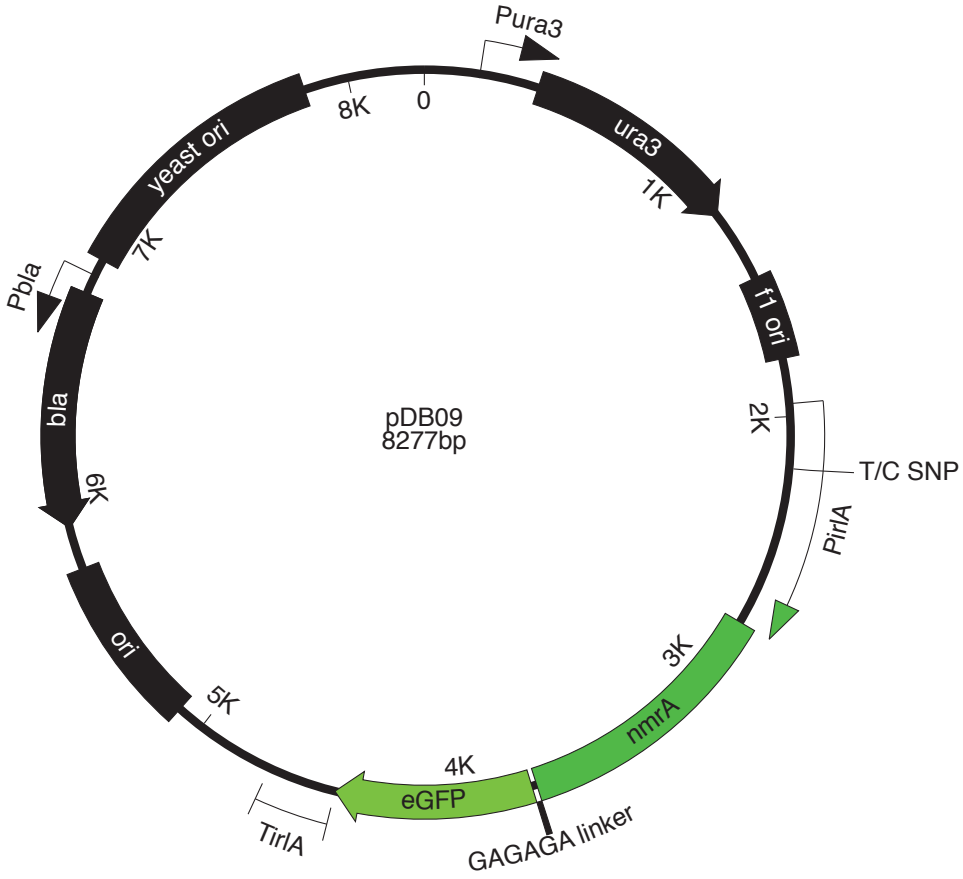
Assembly fragments:

1. 5505 bp vector backbone fragment amplified from linearised pRS426 using primer pair pRS426_F/pRS426_R.
2. 1856 bp P*irlA*–*irlA* fragment encoding 3' GAGAGA linker amplified from *E. festucae* Fl1 gDNA using primer pair pDB08_1/pDB09_1.
3. 754 bp *eGFP* fragment amplified from plasmid containing *eGFP* gene using primer pair pDB09_2/P+GFP_R.
4. 314 bp T*irlA* fragment amplified from *E. festucae* Fl1 gDNA using primer pair pDB09_3/pDB08_2.

Assembly method: Gibson assembly (2.6.12)

Sequencing: T→C SNP in P*irlA* at -584 bp relative to start of *irlA* gene. Was considered highly unlikely to affect transcription initiation.

Plasmid map:



9.3.9 Details of plasmid pDB15

Purpose: Heterologous expression of *perA-ΔR* in *P. paxilli*.

Insert: P*paxM*–*perA-ΔR*–T*paxM*

Vector: pRS426

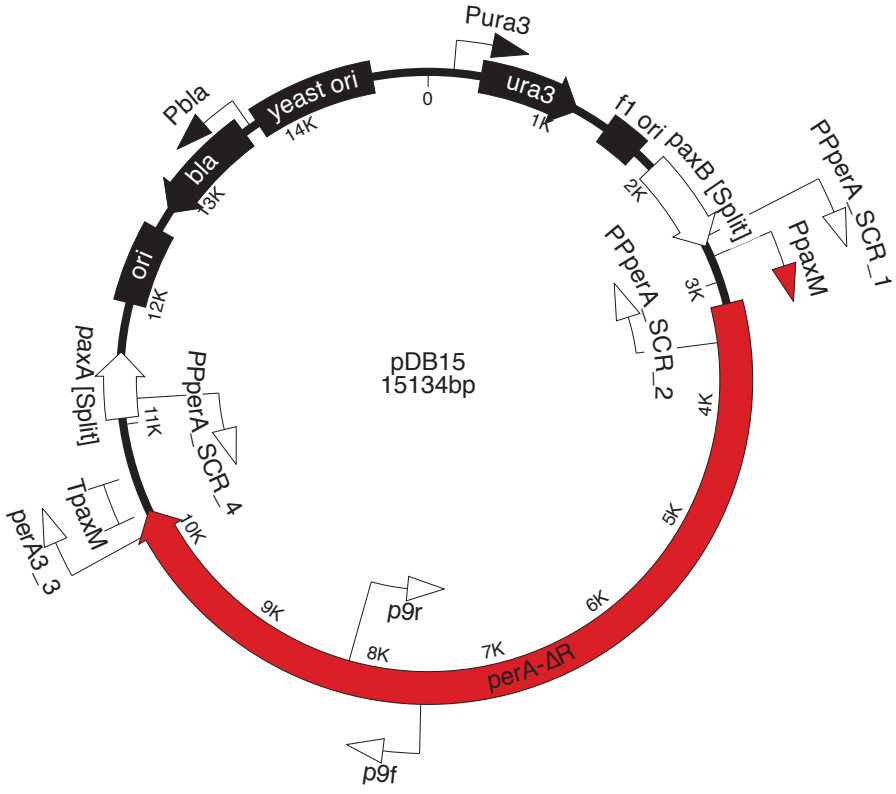
Assembly fragments:

1. 5505 bp vector backbone fragment amplified from linearised pRS426 using primer pair pRS426_F/pRS426_R.
2. 1292 bp P*paxM* fragment PCR amplified from *P. paxilli* gDNA using primer pair pDB05_1/pDB05_2.
3. 2273 bp *perA-ΔR* fragment PCR amplified from *E. festucae* E2368 gDNA using primer pair pDB05_3/pDB15_1.
4. 3019 bp *perA-ΔR* fragment PCR amplified from *E. festucae* E2368 gDNA using primer pair pDB05_5/pDB15_2
5. 1899 bp *perA-ΔR* fragment PCR amplified from *E. festucae* E2368 gDNA using primer pair pDB05_7/pDB15_3.
6. 1307 bp P*paxM* fragment PCR amplified from *P. paxilli* gDNA using primer pair pDB15_4/pDB05_10.

Assembly method: Gibson assembly (2.6.12)

Sequencing: Fidelity of insert sequence confirmed 100%

Plasmid map:



9.3.10 Details of plasmid pDB17

Purpose: Heterologous expression of *perA-17* in *P. paxilli*

Insert: P*paxM–perA-17*–T*paxM*

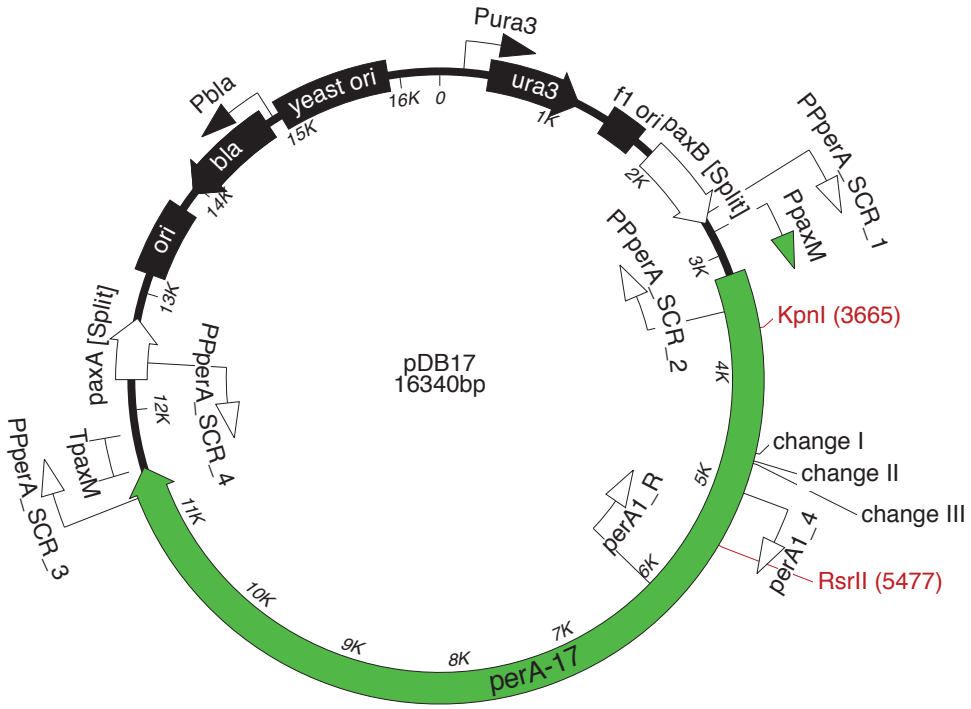
Vector: pDB05

1. 14528 bp vector backbone fragment from digestion of pDB05 with *Kpn*I and *Rsr*II.
2. 1088 bp *perA* A1-domain encoding fragment with introduced *perA-ΔR* SNPs amplified from *E. festucae* Fl1 gDNA with primer pair pDB17_1/pDB17_2.
3. 801 bp *perA* A1-domain encoding fragment with introduced *perA-ΔR* SNPs amplified from *E. festucae* Fl1 gDNA with primer pair pDB17_3/pDB17_4.

Assembly method: Gibson assembly (2.6.12)

Sequencing: Fidelity of insert sequence confirmed 100%

Plasmid map:



9.3.11 Details of plasmid pDB18

Purpose: Heterologous expression of *perA-18* in *P. paxilli*

Insert: P*paxM–perA-18*–T*paxM*

Vector: pDB05

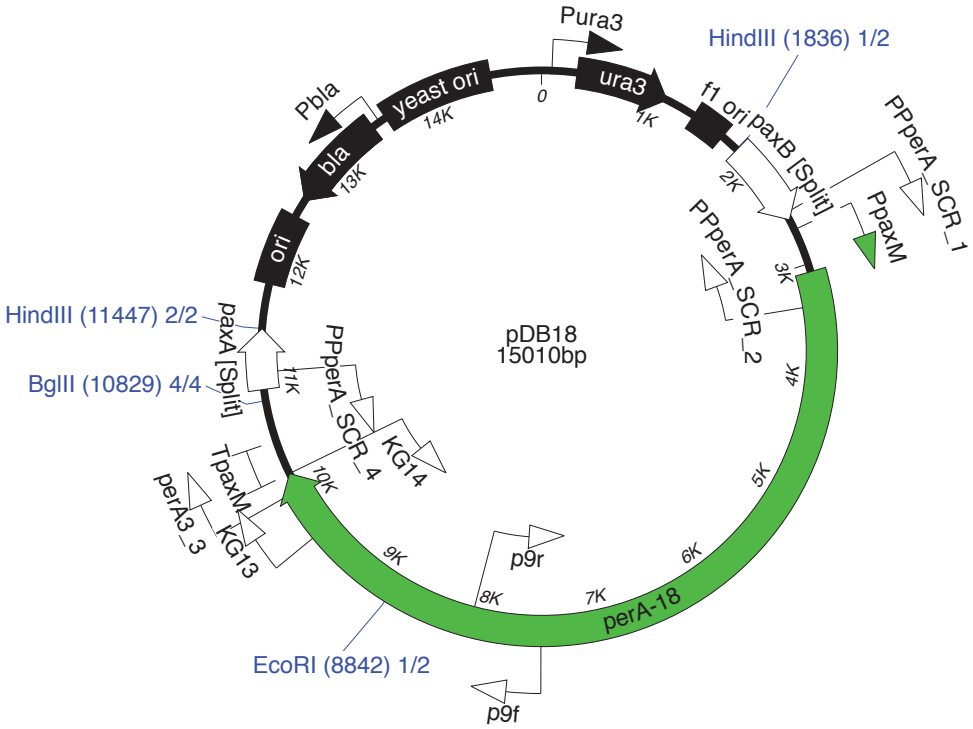
Assembly fragments:

1. 12405 bp vector backbone isolated after complete digestion of pDB05 with *Eco*RI and *Bgl*II, then partial digestion with *Hind*III.
2. 1384 bp *perA* M/T2-domain encoding fragment amplified from *E. festucae* Fl1 gDNA with primer pair pDB18_1/pDB18_2.
3. 1307 bp T*paxM* fragment amplified from *P. paxilli* gDNA with primer pair pDB15_4/pDB05_10.

Assembly method: Gibson assembly (2.6.12)

Sequencing: Fidelity of insert sequence confirmed 100%

Plasmid map (only relevant RE cut sites are shown):



9.3.12 Details of plasmid pDB19

Purpose: Heterologous expression of *perA-19* in *P. paxilli*

Insert: P*paxM*–*perA-19*–T*paxM*

Vector: pDB05

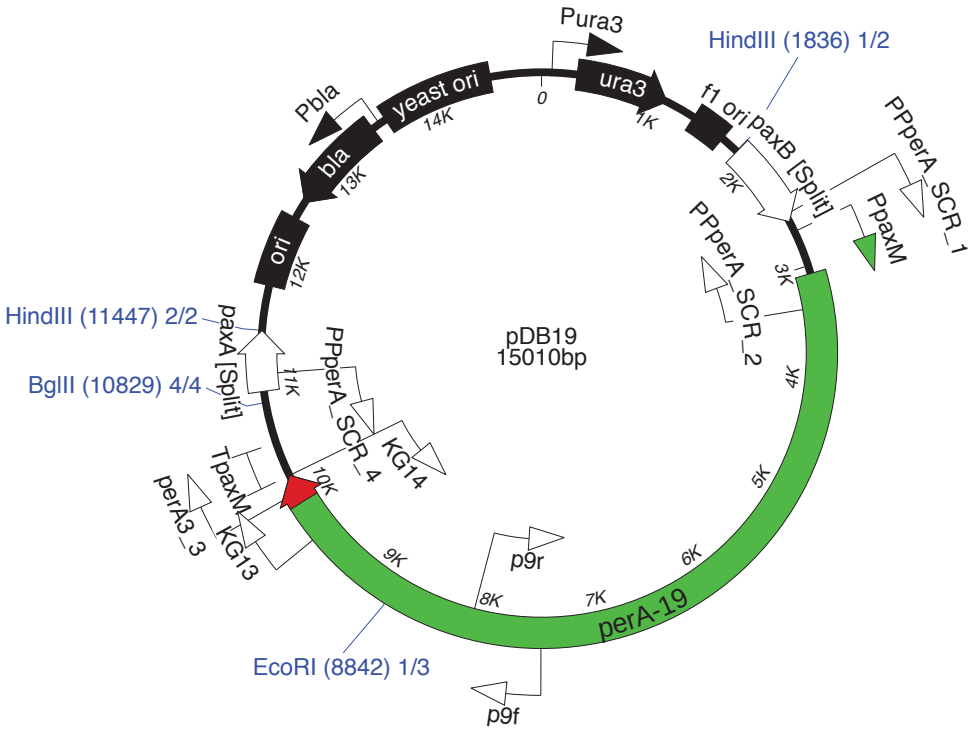
Assembly fragments:

1. 12405 bp vector backbone isolated after complete digestion of pDB05 with *Eco*RI and *Bgl*II, then partial digestion with *Hind*III.
2. 1107 bp *perA* M-domain encoding fragment amplified from *E. festucae* Fl1 gDNA with primer pair pDB18_1/pDB19_1.
3. 305 bp *perA-ΔR* T2-domain encoding fragment amplified from *E. festucae* E2368 gDNA with primer pair pDB19_2/pDB15_3.
4. 1307 bp T*paxM* fragment amplified from *P. paxilli* gDNA with primer pair pDB15_4/pDB05_10.

Assembly method: Gibson assembly (2.6.12)

Sequencing: Fidelity of insert sequence confirmed 100%

Plasmid map (only relevant RE cut sites are shown):



9.3.13 Details of plasmid pDB20

Purpose: Heterologous expression of *perA-20* in *P. paxilli*

Insert: P*paxM*–*perA-20*–T*paxM*

Vector: pRS426

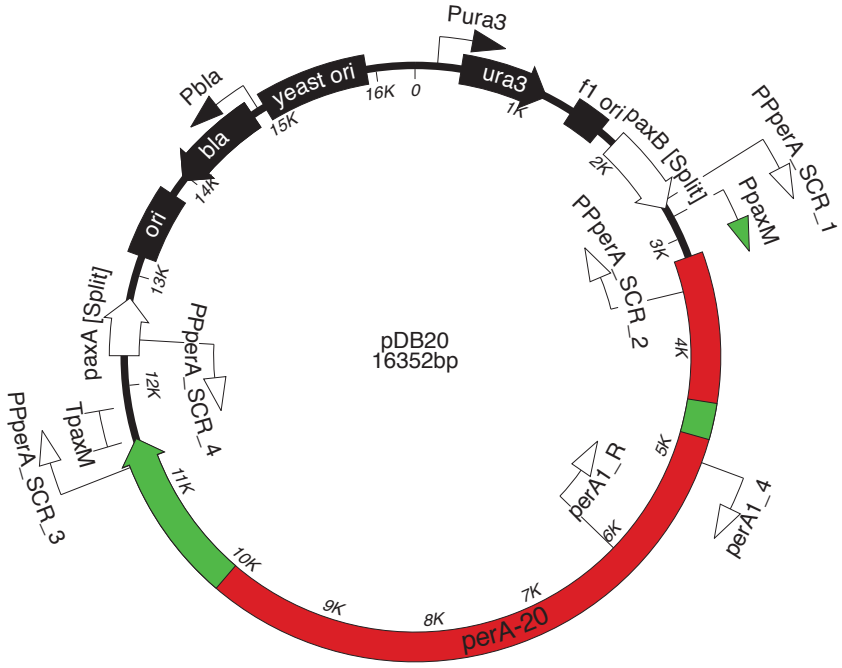
Assembly fragments:

1. 5505 bp vector backbone fragment amplified from linearised pRS426 using primer pair pRS426_F/pRS426_R.
2. 1292 bp P*paxM* fragment PCR amplified from *P. paxilli* gDNA using primer pair pDB05_1/pDB05_2.
3. 1339 bp *perA-ΔR* fragment amplified from *E. festucae* E2368 gDNA using primer pair pDB05_3/pDB20_1.
4. 315 bp *perA* A1-domain substrate-binding region encoding fragment amplified from *E. festucae* Fl1 gDNA using primer pair pDB20_2/pDB20_3.
5. 5269 bp *perA-ΔR* fragment amplified from *E. festucae* E2368 gDNA using primer pair pDB20_4/pDB19_1.
6. 1497 bp *perA* R-domain encoding fragment amplified from *E. festucae* Fl1 gDNA using primer pair pDB19_2/pDB05_8.
7. 1336 bp P*paxM* fragment PCR amplified from *P. paxilli* gDNA using primer pair pDB05_9/pDB05_10.

Assembly method: Gibson assembly (2.6.12)

Sequencing: Fidelity of insert sequence confirmed 100%

Plasmid map:



9.3.14 Details of plasmid pDB24

Purpose: *pdtA* gene deletion construct

Insert: *pdtA* LB–P*trpC–hph–pdtA* RB

Vector: pRS426

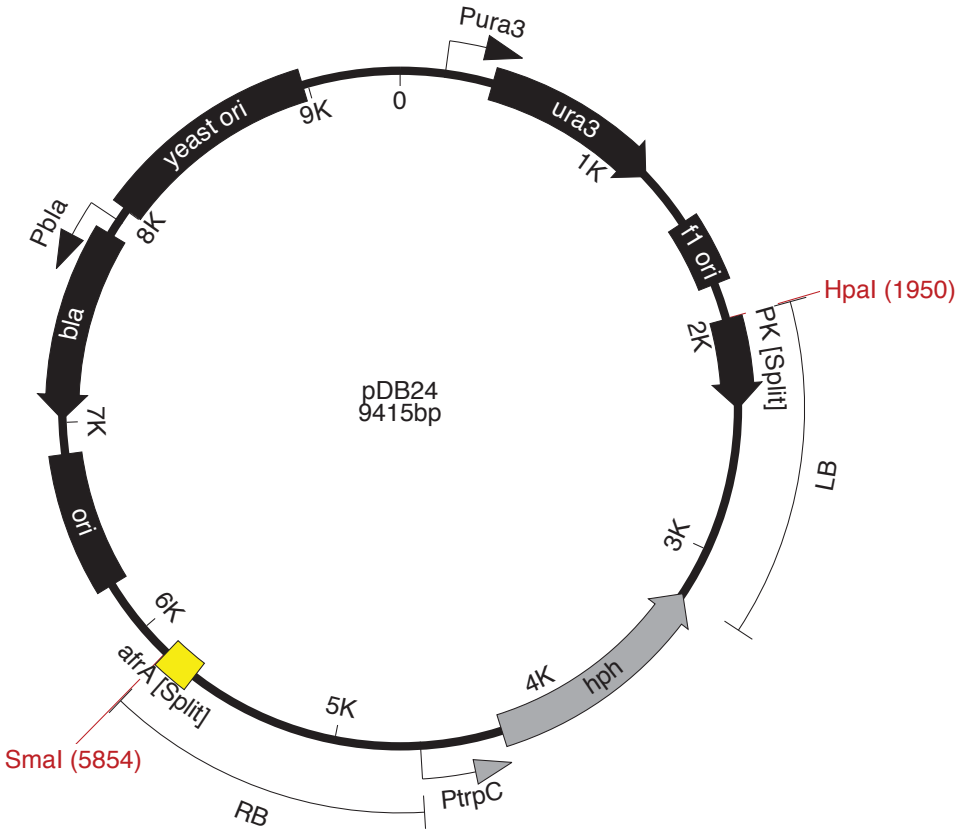
Assembly fragments:

1. 5505 bp vector backbone fragment amplified from linearised pRS426 using primer pair pRS426_F/pRS426_R.
2. 1326 bp *pdtA* LB fragment PCR amplified from *E. festucae* Fl1 gDNA using primer pair pDB24_1/pDB24_2.
3. 1394 bp P*trpC-hph* fragment PCR amplified from pSF15.15 using primer pair hyg_F/hyg_R.
4. 1290 bp *pdtA* RB fragment PCR amplified from *E. festucae* Fl1 gDNA using primer pair pDB24_3/pDB24_4.

Assembly method: Gibson assembly (2.6.12)

Sequencing: Possible 1×A deletion within 10×A mononucleotide run located in non-coding region of right border. Considered unlikely to cause complications.

Plasmid map:



9.3.15 Details of plasmid pDB25

Purpose: *afrA* gene deletion construct

Insert: *afrA* LB–P*trpC–hph–afrA* RB

Vector: pRS426

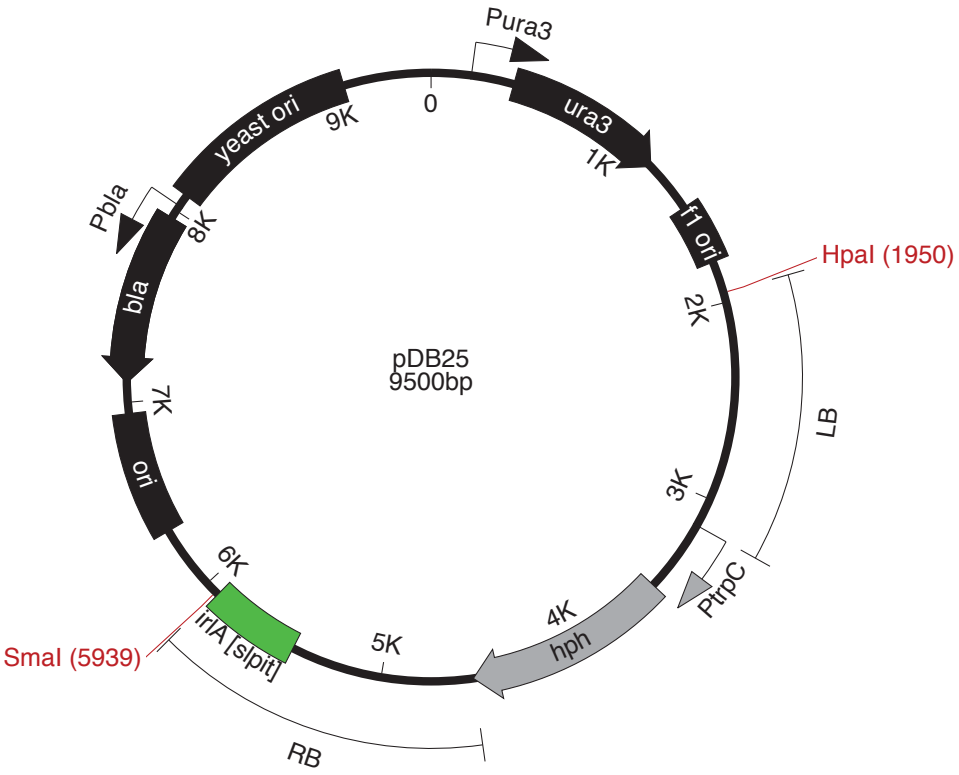
Assembly fragments:

1. 5505 bp vector backbone fragment amplified from linearised pRS426 using primer pair pRS426_F/pRS426_R.
2. 1248 bp *afrA* LB fragment PCR amplified from *E. festucae* Fl1 gDNA using primer pair pDB25_1/pDB25_2.
3. 1394 bp P*trpC-hph* fragment PCR amplified from pSF15.15 using primer pair hyg_F/hyg_R.
4. 1453 bp *afrA* RB fragment PCR amplified from *E. festucae* Fl1 gDNA using primer pair pDB25_3/pDB25_4.

Assembly method: Gibson assembly (2.6.12)

Sequencing: Fidelity of insert sequence confirmed 100%

Plasmid map:



9.3.16 Details of plasmid pDB26

Purpose: *mfsB* gene deletion construct

Insert: *mfsB* LB–P*trpC–hph–mfsB* RB

Vector: pRS426

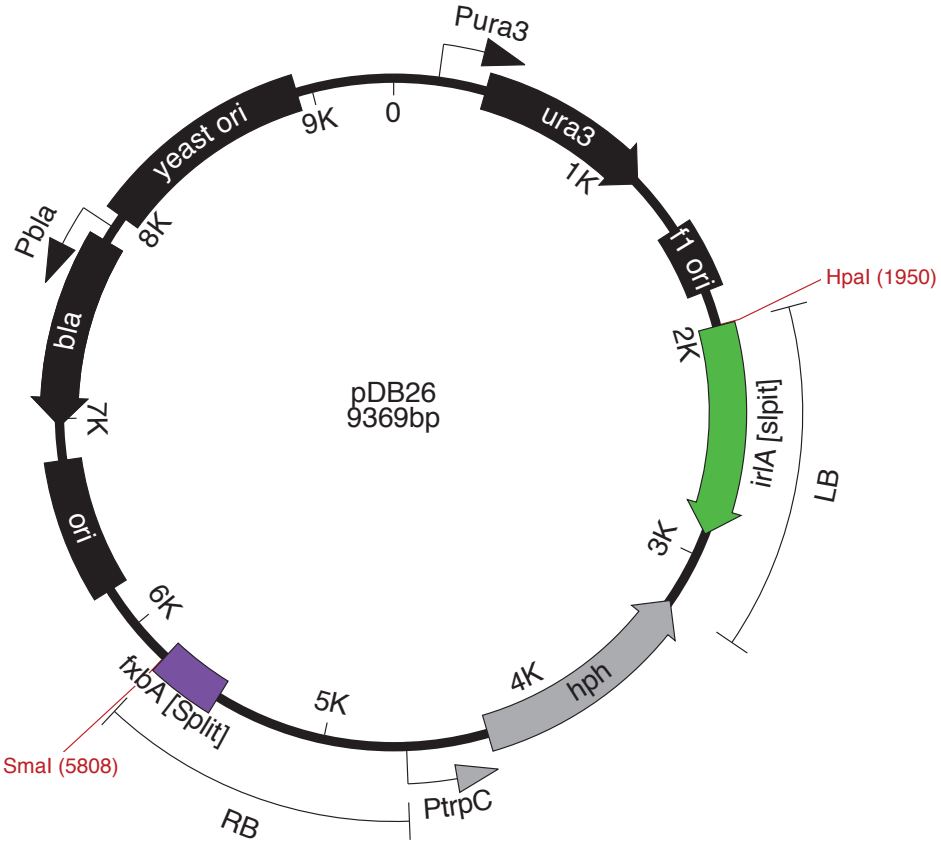
Assembly fragments:

1. 5505 bp vector backbone fragment amplified from linearised pRS426 using primer pair pRS426_F/pRS426_R.
2. 1336 bp *mfsB* LB fragment PCR amplified from *E. festucae* Fl1 gDNA using primer pair pDB26_1/pDB26_2.
3. 1394 bp P*trpC-hph* fragment PCR amplified from pSF15.15 using primer pair hyg_F/hyg_R.
4. 1234 bp *mfsB* RB fragment PCR amplified from *E. festucae* Fl1 gDNA using primer pair pDB26_3/pDB26_4.

Assembly method: Gibson assembly (2.6.12)

Sequencing: Fidelity of insert sequence confirmed 100%

Plasmid map:



9.3.17 Details of plasmid pDB27

Purpose: *fxbA* gene deletion construct

Insert: *fxbA* LB–P*trpC–hph–fxbA* RB

Vector: pRS426

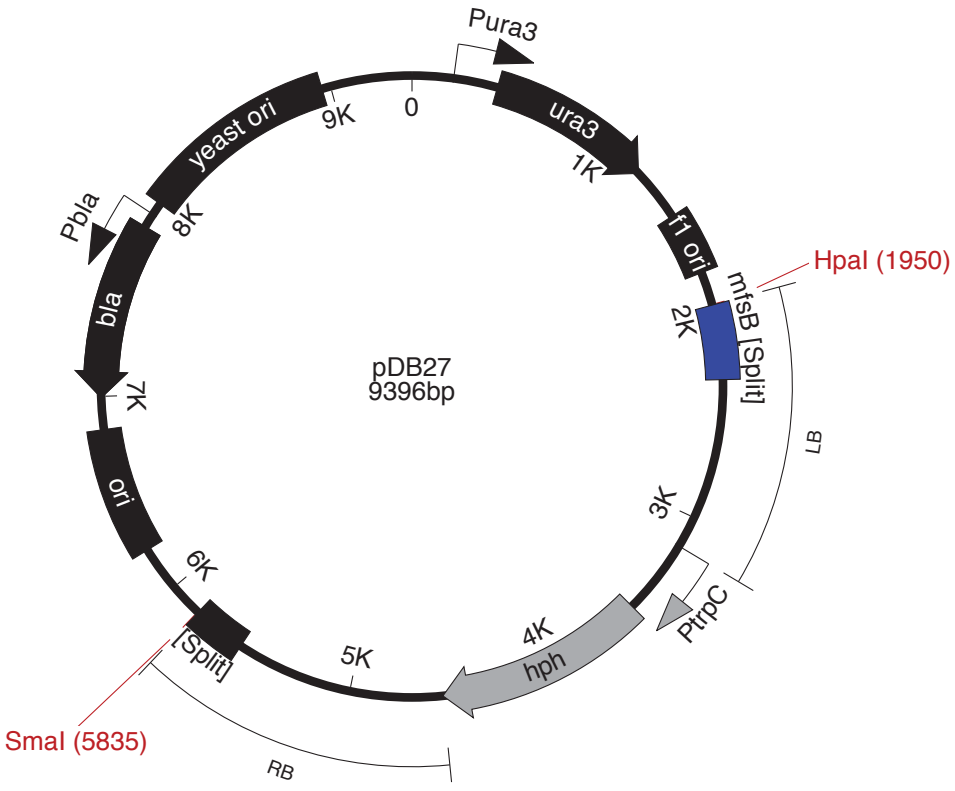
Assembly fragments:

1. 5505 bp vector backbone fragment amplified from linearised pRS426 using primer pair pRS426_F/pRS426_R.
2. 1258 bp *fxbA* LB fragment PCR amplified from *E. festucae* Fl1 gDNA using primer pair pDB27_1/pDB27_2.
3. 1394 bp P*trpC-hph* fragment PCR amplified from pSF15.15 using primer pair hyg_F/hyg_R.
4. 1339 bp *fxbA* RB fragment PCR amplified from *E. festucae* Fl1 gDNA using primer pair pDB27_3/pDB27_4.

Assembly method: Gibson assembly (2.6.12)

Sequencing: Fidelity of insert sequence confirmed 100%

Plasmid map:



9.3.18 Details of plasmid pDB29

Purpose: Heterologous expression of *perA-29* in *P. paxilli*

Insert: P*paxM–perA-29*–T*paxM*

Vector: pDB20

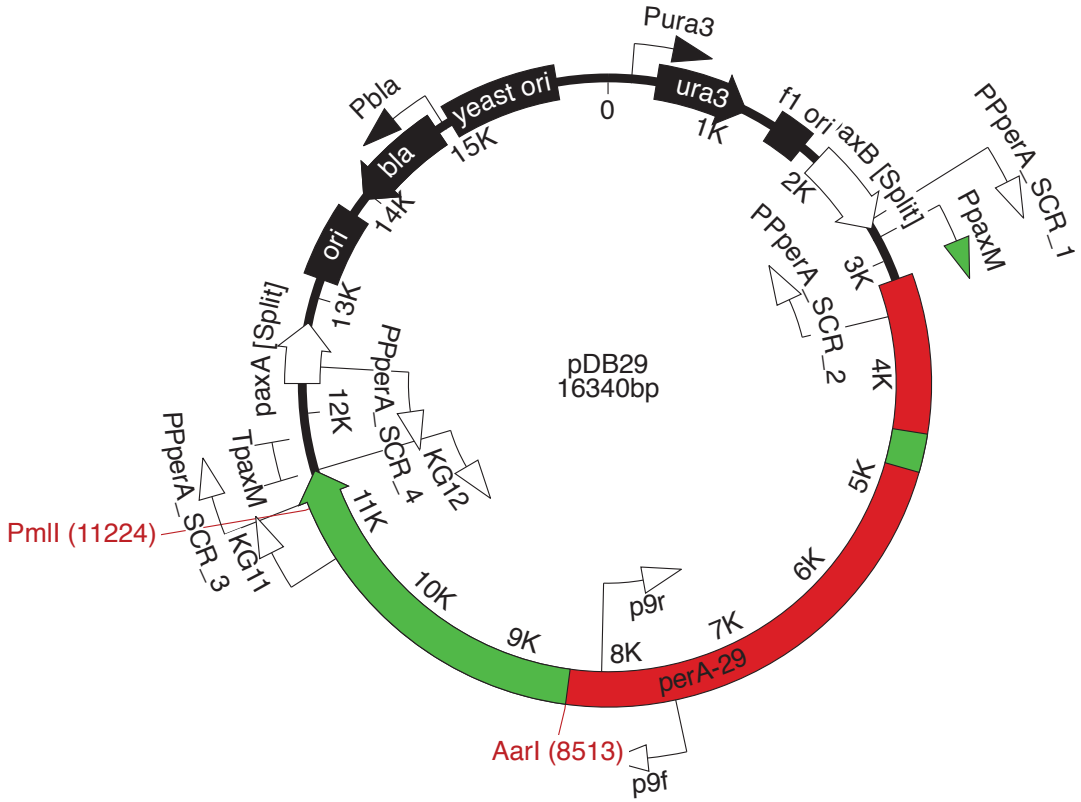
Assembly fragments:

1. 13629 bp vector backbone fragment isolated after digestion of pDB20 with *Pml*I and *Aar*I.
2. 2766 bp *perA* 3' fragment PCR amplified from *E. festucae* Fl1 gDNA using primer pair pDB29_1/pDB29_2.

Assembly method: Gibson assembly (2.6.12)

Sequencing: Fidelity of insert sequence confirmed 100%

Plasmid map:



9.3.19 Details of plasmid pDB30

Purpose: Heterologous expression of *perA-30* in *P. paxilli*

Insert: P*paxM–perA-30*–T*paxM*

Vector: pRS426

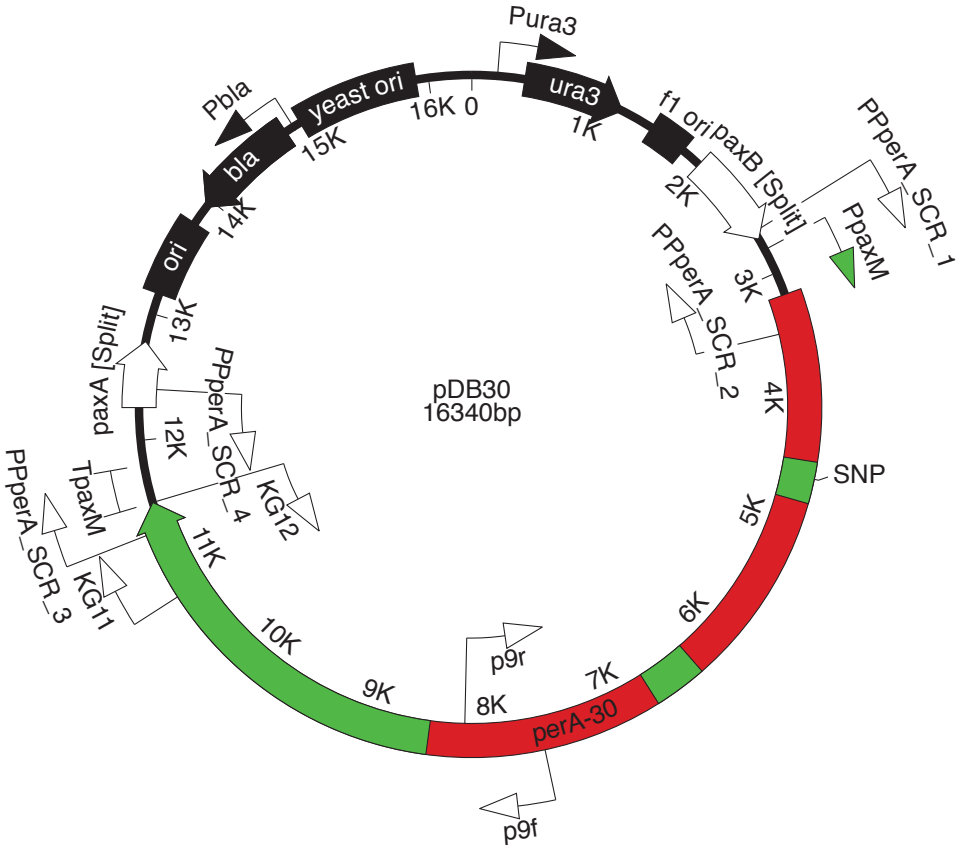
Assembly fragments:

1. 5505 bp vector backbone fragment amplified from linearised pRS426 using primer pair pRS426_F/pRS426_R.
2. 4409 bp *perA-29* fragment PCR amplified from linear pDB29 using primer pair pDB05_1/pDB30_5.
3. 403 bp *perA-ΔR* C-domain encoding fragment amplified from *E. festucae* E2368 gDNA using primer pair pDB30_2/pDB30_3.
4. 6130 bp *perA-29* fragment amplified from linear pDB29 using primer pair pDB30_6/pDB05_10.

Assembly method: Gibson assembly (2.6.12)

Sequencing: Contains a synonymous C→A SNP at position 1440 relative to start of *perA-30*.

Plasmid map:



9.3.20 Details of plasmid pDB31

Purpose: Heterologous expression of *perA-31* in *P. paxilli*

Insert: P*paxM–perA-31*–T*paxM*

Vector: pDB05

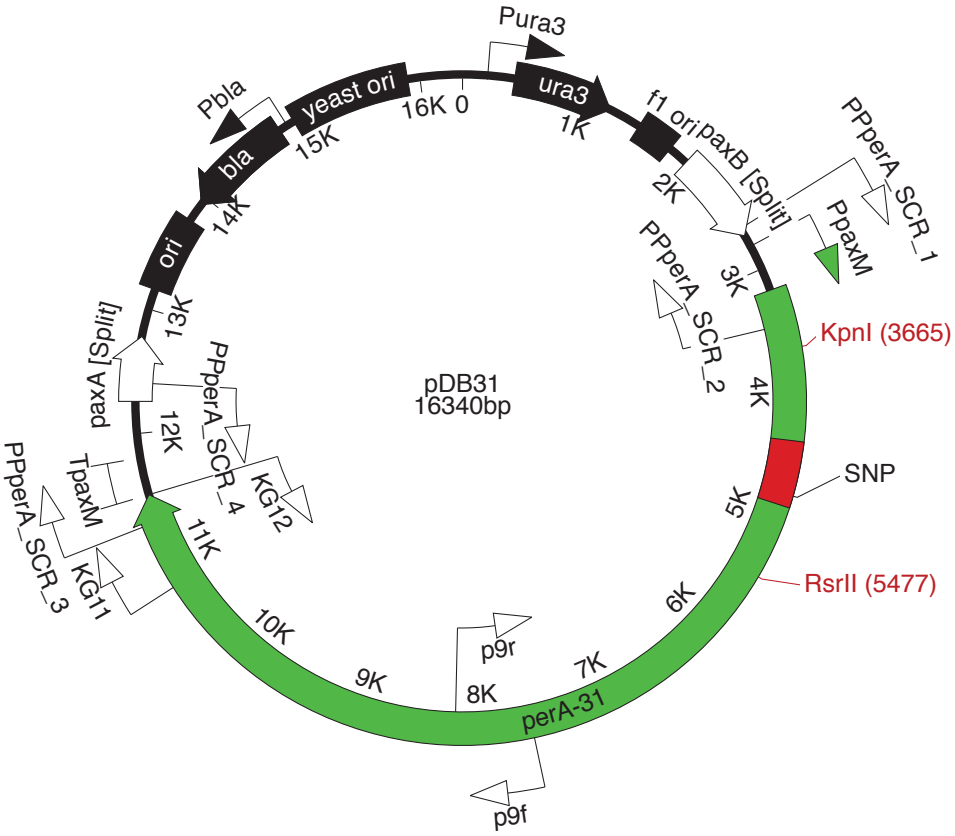
Assembly fragments:

5. 14528 bp vector backbone fragment isolated after digestion of pDB05 with *Kpn*I and *Rsr*II.
6. 795 bp *perA* fragment PCR amplified from *E. festucae* Fl1 gDNA using primer pair pDB17_1/pDB31_1.
7. 512 bp *perA-ΔR* A1-domain substrate-binding region encoding fragment amplified from *E. festucae* E2368 gDNA using primer pair pDB31_2/pDB31_3.
8. 652 bp *perA* fragment amplified from *E. festucae* Fl1 gDNA using primer pair pDB31_4/pDB17_4.

Assembly method: Gibson assembly (2.6.12)

Sequencing: Contains a synonymous G→T SNP at position 1629 relative to start of *perA-31.*

Plasmid map:



9.3.21 Details of plasmid pDB32

Purpose: Heterologous expression of *perA-32* in *P. paxilli*

Insert: P*paxM–perA-32*–T*paxM*

Vector: pRS426

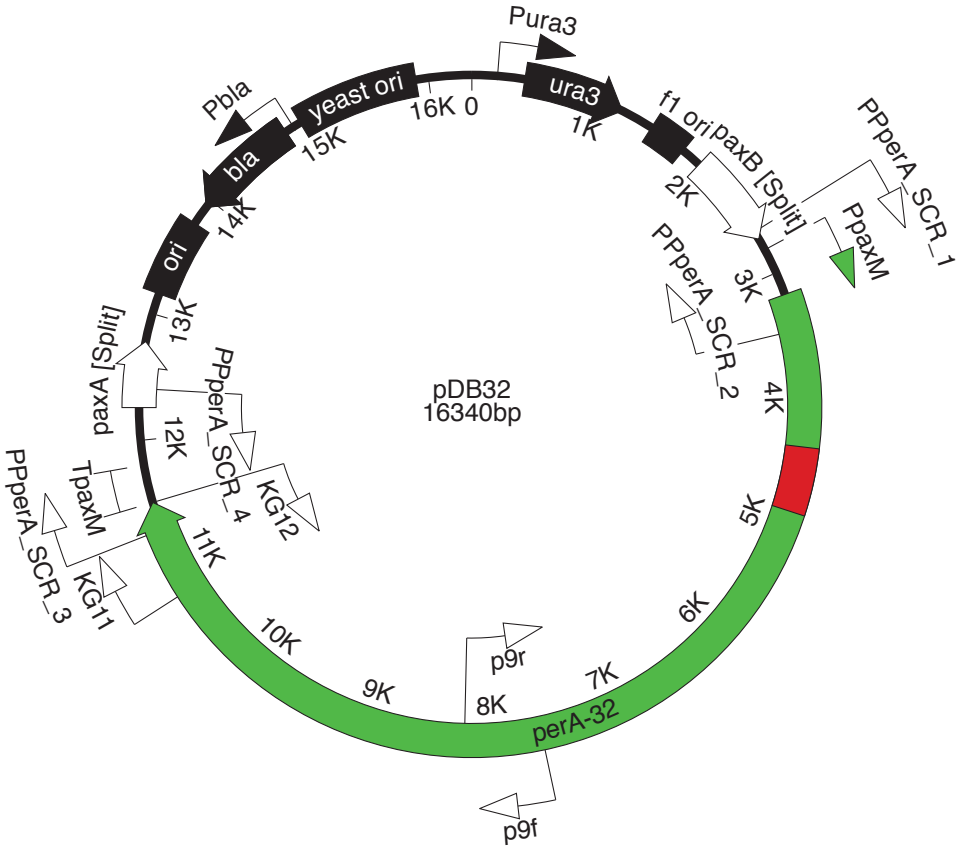
Assembly fragments:

1. 5505 bp vector backbone fragment amplified from linearised pRS426 using primer pair pRS426_F/pRS426_R.
2. 2499 bp *perA* fragment PCR amplified from linear pDB05 using primer pair pDB05_1/pDB31_1.
3. 512 bp *perA-ΔR* A1-domain substrate-binding region encoding fragment amplified from *E. typhina* ssp. *poae* gDNA using primer pair pDB31_2/pDB31_3.
4. 7952 bp *perA* fragment amplified from linearised pDB05 using primer pair pDB31_4/pDB05_10.

Assembly method: Gibson assembly (2.6.12)

Sequencing: Fidelity of insert sequence confirmed 100%

Plasmid map:



9.3.22 Details of plasmid pDB34

Purpose: Strong heterologous expression of *E. festucae* E189 *tesB* in *P. paxilli*

Insert: P*tefA*– *tesB*(*Efes*)–T*tesB*(*Ppax*); P*trpC*–*hph*–T*trpC*

Vector: pRS426

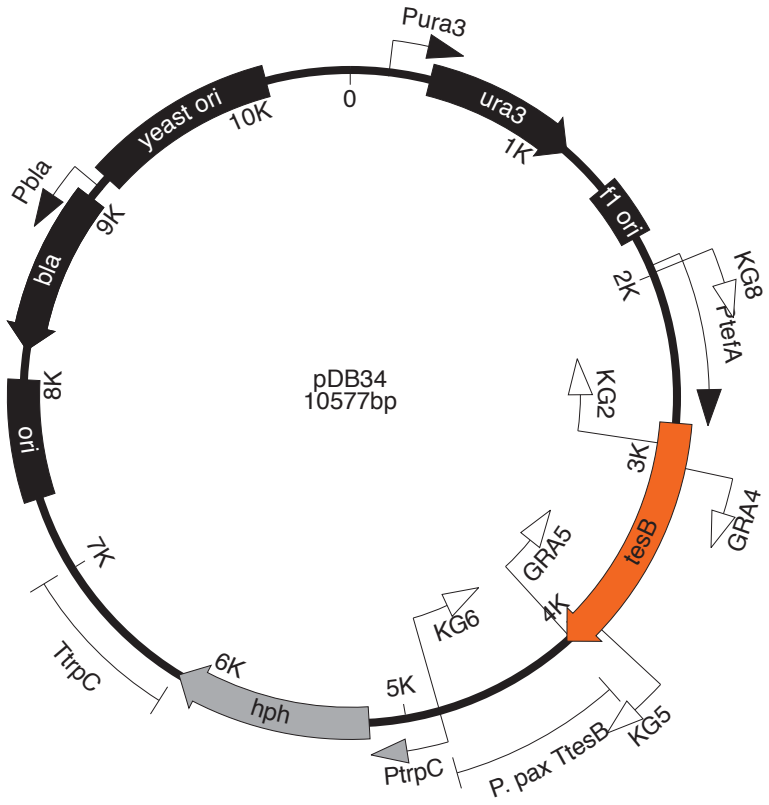
Assembly fragments:

1. 5505 bp vector backbone fragment amplified from linearised pRS426 using primer pair pRS426_F/pRS426_R.
2. 884 bp P*tefA* fragment amplified from a P*tefA*–*eGFP* plasmid using primer pair pDB34_1/pDB34_2..
3. 1309 bp *tesB* fragment amplified from *E. festucae* E189 gDNA using primer pair pDB34_3/pDB34_4.
4. 780 bp T*tesB* fragment amplified from *P. paxilli* gDNA using primer pair pDB34_5/pDB34_6.
5. 1419 bp P*trpC*–*hph* fragment amplified from pSF15.15 using primer pair hyg_F/pDB28_6.
6. 822 bp T*trpC* fragment amplified from pAN7-1 using primer pair pDB28_7/pDB28_8.

Assembly method: Gibson assembly (2.6.12)

Sequencing: Fidelity of insert sequence confirmed 100%

Plasmid map:



9.3.23 Details of plasmid pDB35

Purpose: Heterologous expression of *E. festucae* E189 *tesB* in *P. paxilli*

Insert: P*tesB*(*Ppax*)–*tesB*(*Efes*)–T*tesB*(*Ppax*); P*trpC*–*hph*–T*trpC*

Vector: pRS426

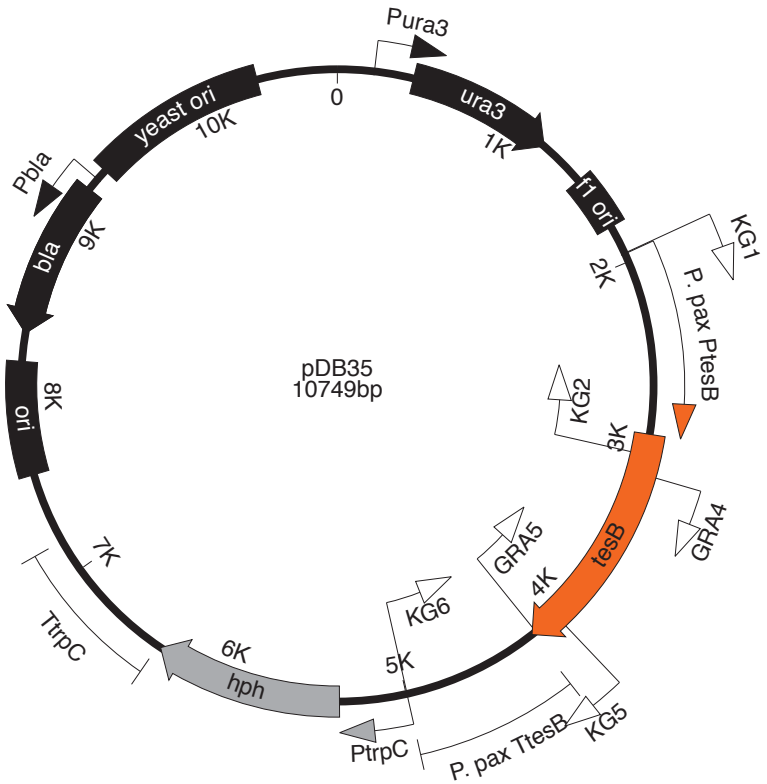
Assembly fragments:

1. 5505 bp vector backbone fragment amplified from linearised pRS426 using primer pair pRS426_F/pRS426_R.
2. 1056 bp P*tesB* fragment amplified from *P. paxilli* gDNA using primer pair pDB35_1/pDB35_2.
3. 1309 bp *tesB* fragment amplified from *E. festucae* E189 gDNA using primer pair pDB34_3/pDB34_4.
4. 780 bp T*tesB* fragment amplified from *P. paxilli* gDNA using primer pair pDB34_5/pDB34_6.
5. 1419 bp P*trpC–hph* fragment amplified from pSF15.15 using primer pair hyg_F/pDB28_6.
6. 822 bp T*trpC* fragment amplified from pAN7-1 using primer pair pDB28_7/pDB28_8.

Assembly method: Gibson assembly (2.6.12)

Sequencing: Fidelity of insert sequence confirmed 100%

Plasmid map:



9.3.24 Details of plasmid pDB36

Purpose: *irlA* deletion construct

Insert: *irlA* LB–P*trpC–hph–irlA* RB

Vector: pRS426.

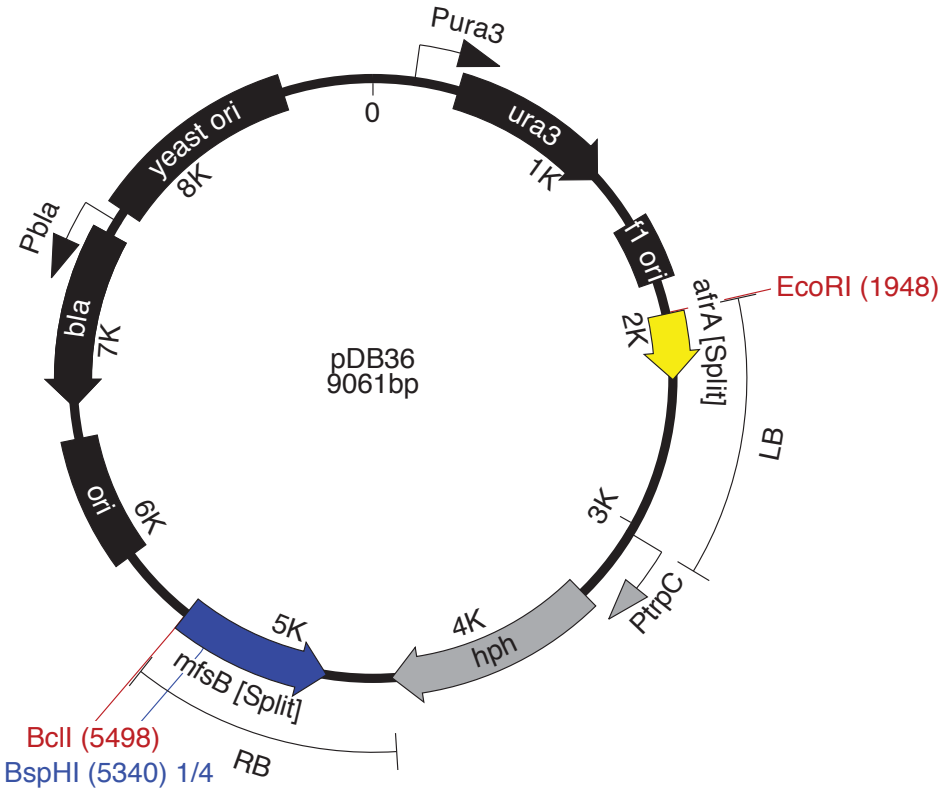
Assembly fragments:

7. 5505 bp vector backbone fragment amplified from linearised pRS426 using primer pair pRS426_F/pRS426_R.
8. 1154 bp *irlA* LB fragment PCR amplified from *E. festucae* Fl1 gDNA using primer pair pDB02_1/pDB02_2.
9. 1394 bp P*trpC-hph* fragment PCR amplified from pSF15.15 using primer pair hyg_F/hyg_R.
10. 1122 bp *irlA* RB fragment PCR amplified from *E. festucae* Fl1 gDNA using primer pair pDB02_3/pDB36_1.

Assembly method: Gibson assembly (2.6.12)

Sequencing: Fidelity of insert sequence confirmed 100%

Plasmid map:



9.3.25 Details of plasmid pDB43

Purpose: Heterologous expression of *perA-43* in *P. paxilli*

Insert: P*paxM–perA-43*–T*paxM*

Vector: pRS426

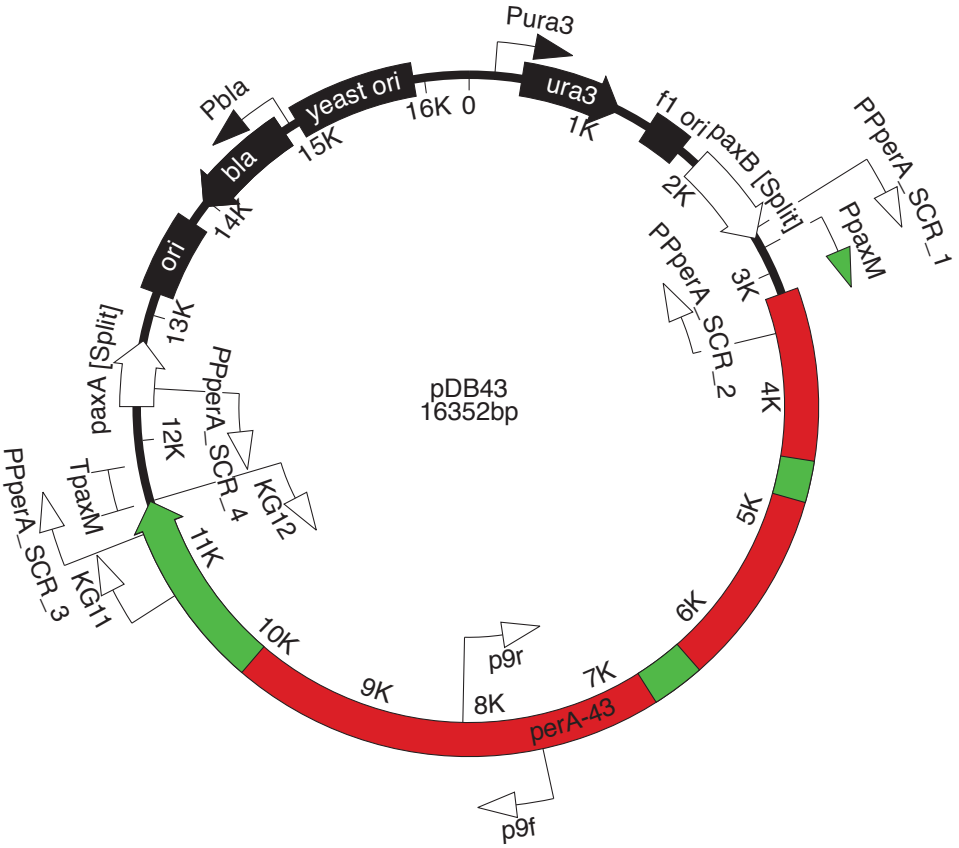
Assembly fragments:

9. 5505 bp vector backbone fragment amplified from linearised pRS426 using primer pair pRS426_F/pRS426_R.
10. 4409 bp *perA-20* fragment PCR amplified from linear pDB20 using primer pair pDB05_1/pDB30_5.
11. 403 bp *perA-ΔR* C-domain encoding fragment amplified from *E. festucae* E2368 gDNA using primer pair pDB30_2/pDB30_3.
12. 6142 bp *perA-20* fragment amplified from linear pDB20 using primer pair pDB30_6/pDB05_10.

Assembly method: Gibson assembly (2.6.12)

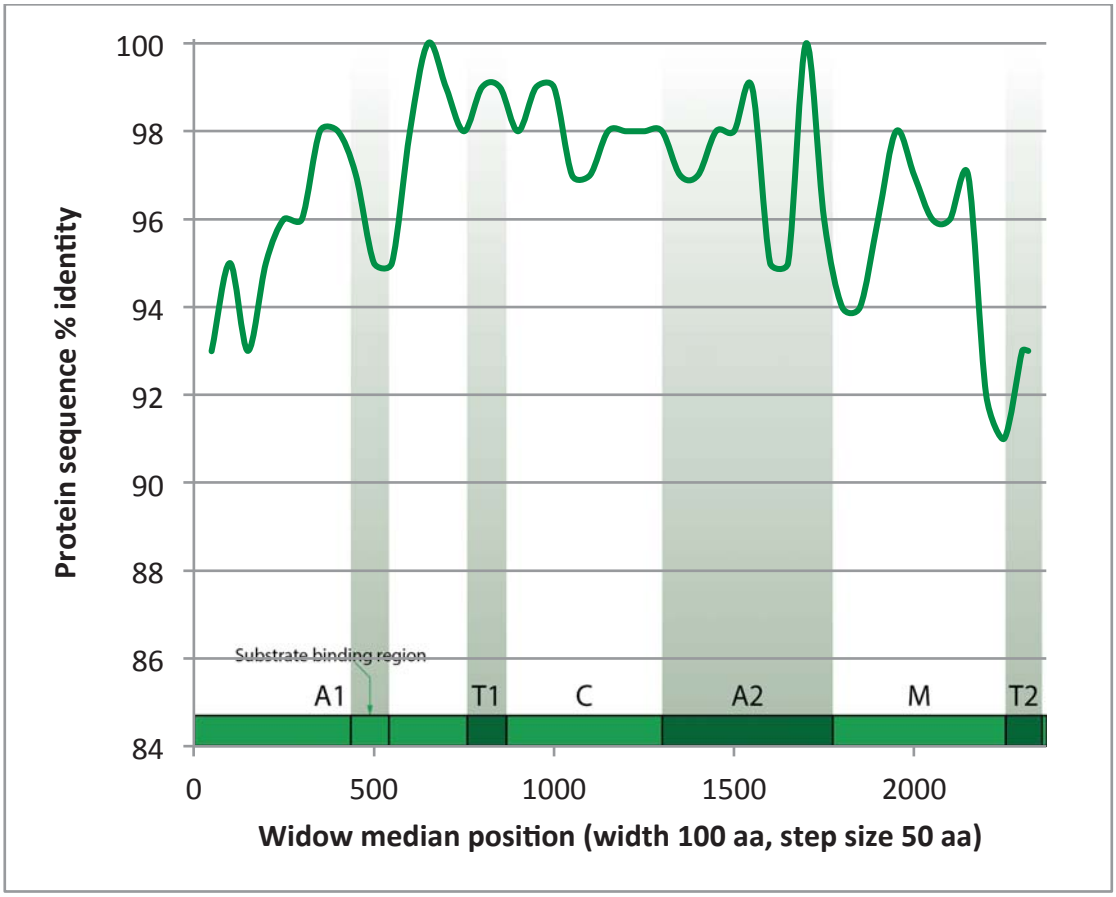
Sequencing: Fidelity of insert sequence confirmed 100%

Plasmid map:

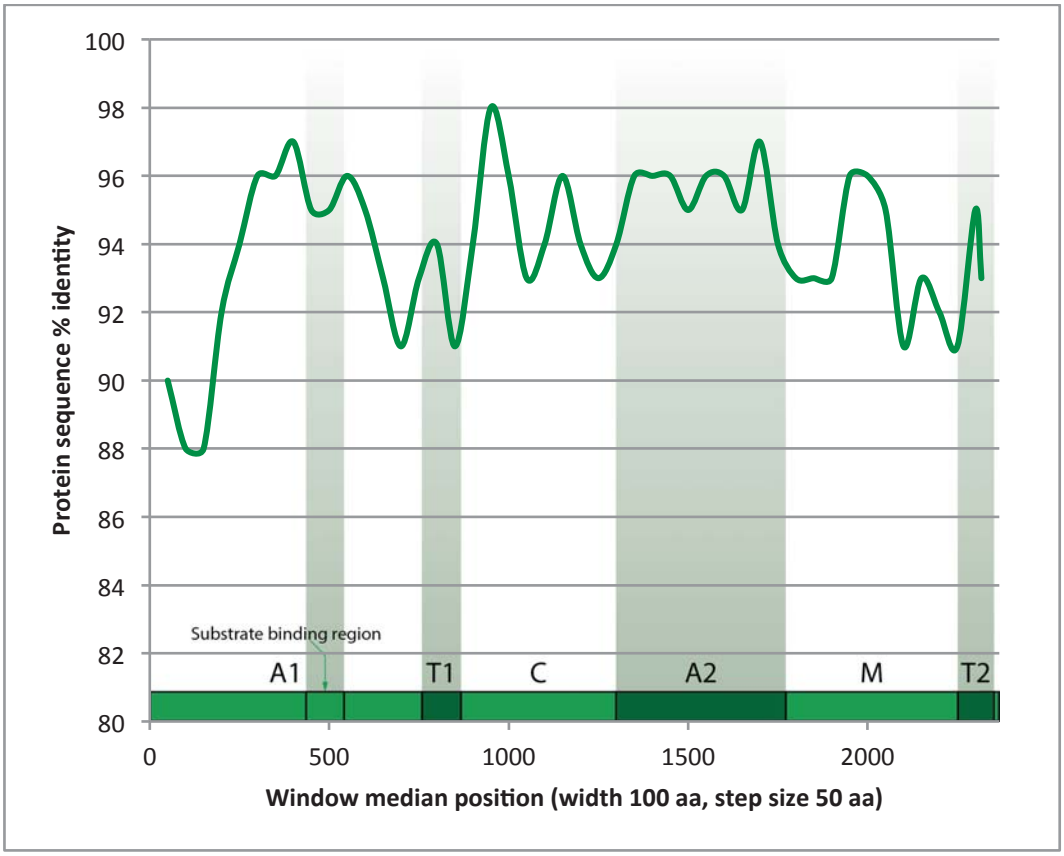


9.4 Supplementary graphs

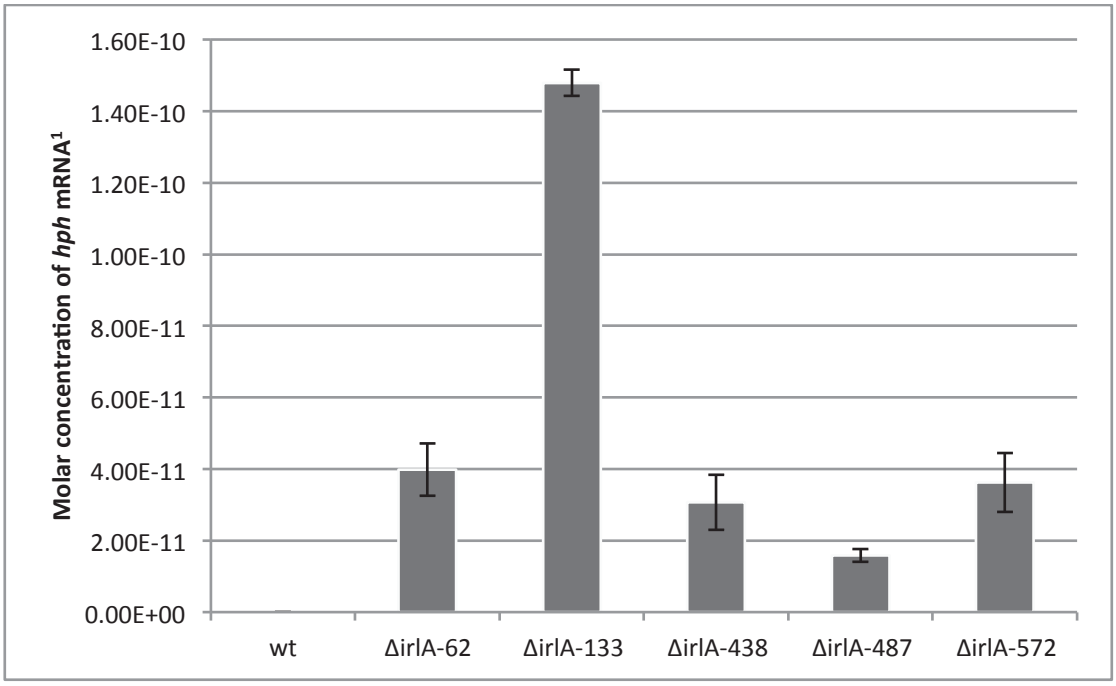
9.4.1 Sliding window protein sequence comparison between *E. festucae* Fl1 PerA and *E. amarillans* E57 PerA



9.4.2 Sliding window protein sequence comparison between *E. festucae* Fl1 PerA and *E. typhina* E8 PerA

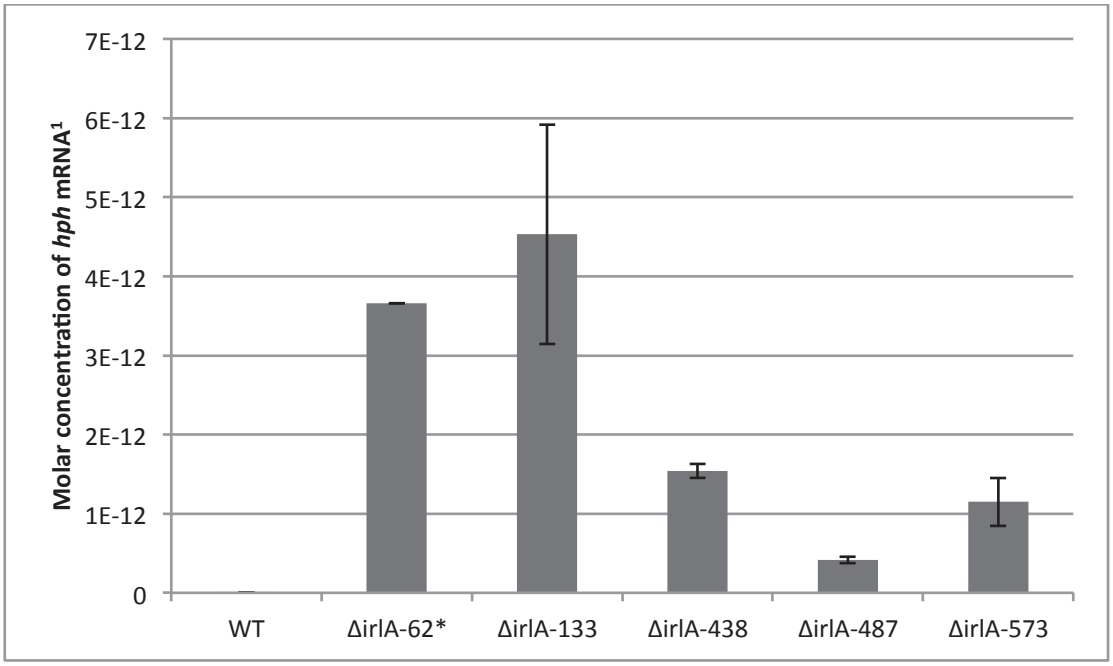


9.4.3 RT-qPCR analysis of *hph* expression for *irlA* mutants in culture



[1] Calculated molar concentration of *hph* mRNA normalised against *EF2* expression.

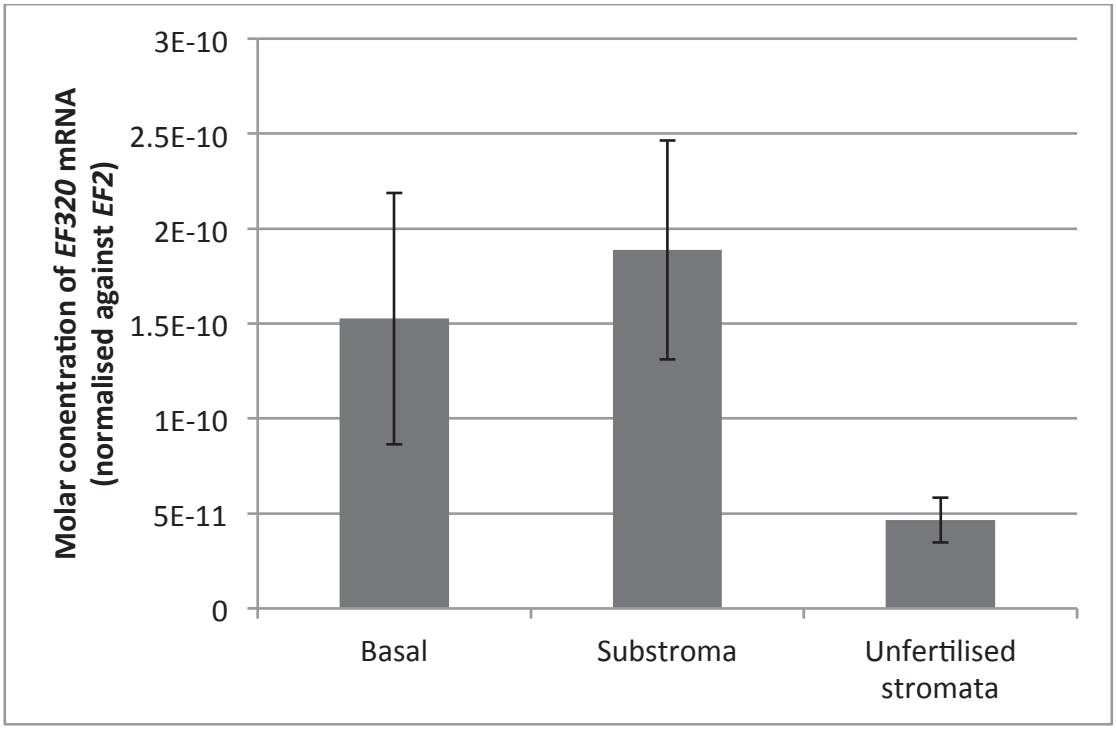
9.4.4 RT-qPCR analysis of *hph* expression for *irlA* mutants *in planta*



[1] Calculated molar concentration of *hph* mRNA normalised against *EF2* expression.

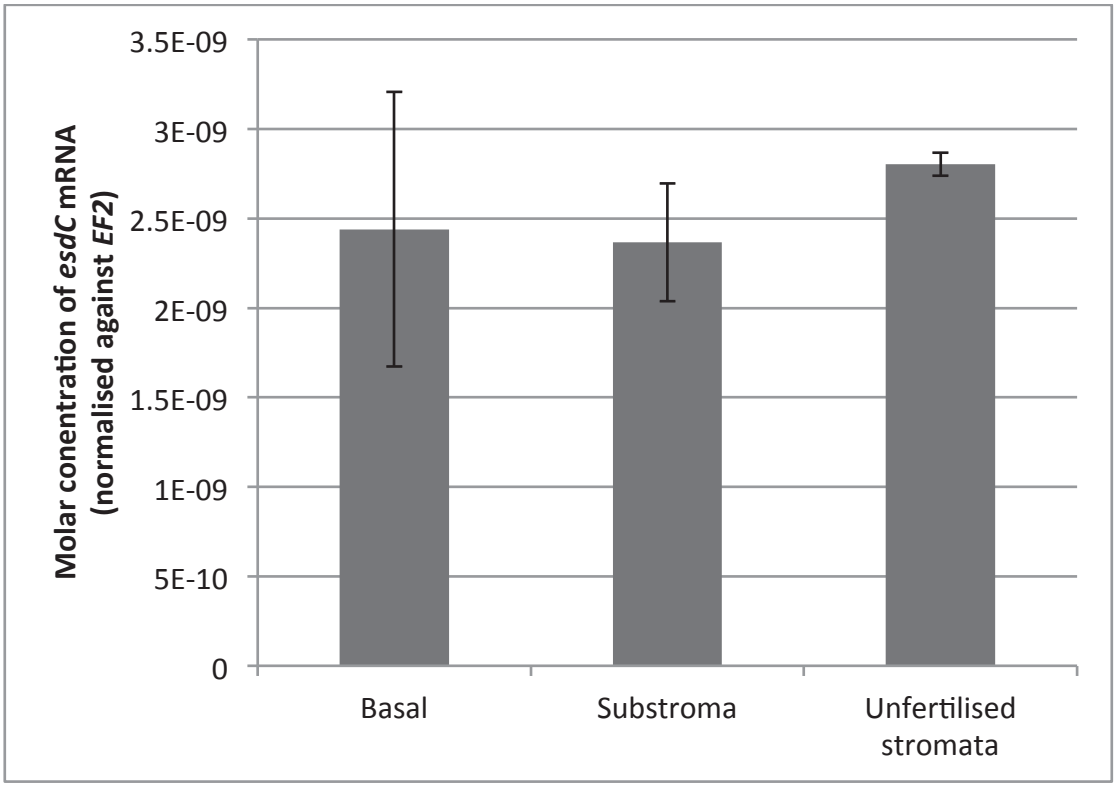
* Error bars not displayed for Δ*irlA*-62 due to lack of replicates.

9.4.5 RT-qPCR analysis of *EF320* expression in *E. typhina*-infected *D. glomerata* tillers exhibiting unfertilised stromata[1]



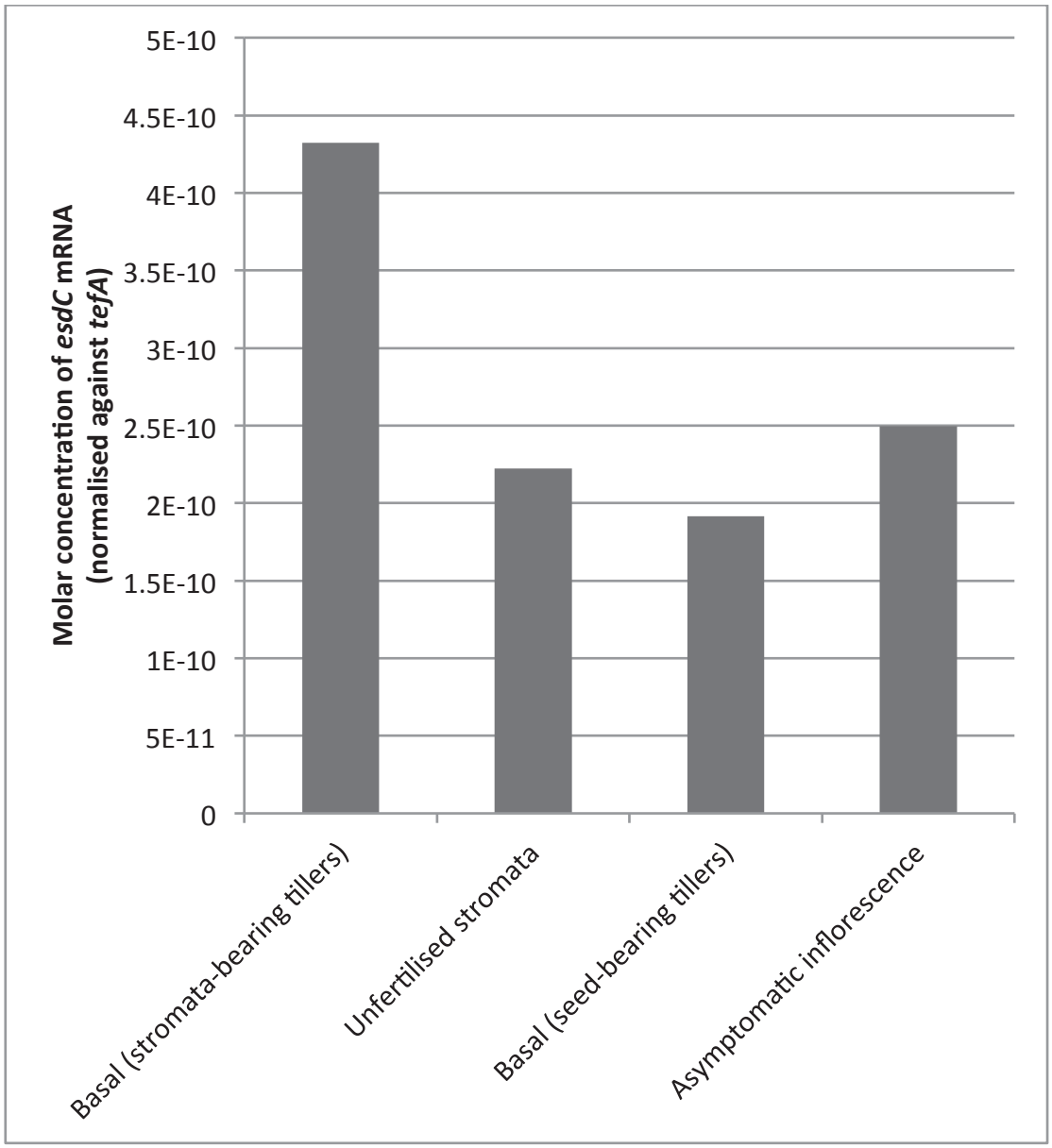
[1] Error bars show standard error of mean

9.4.1 RT-qPCR analysis of *esdC* expression in *E. typhina*-infected *D. glomerata* tillers exhibiting unfertilised stromata[1]



[1] Error bars show standard error of mean

9.4.1 RT-qPCR analysis of *esdC* expression in *E. elymi*-infected *Elymus* sp. tillers exhibiting unfertilised stromata and asymptomatic inflorescences[1]



[1] Single biological replicates only.

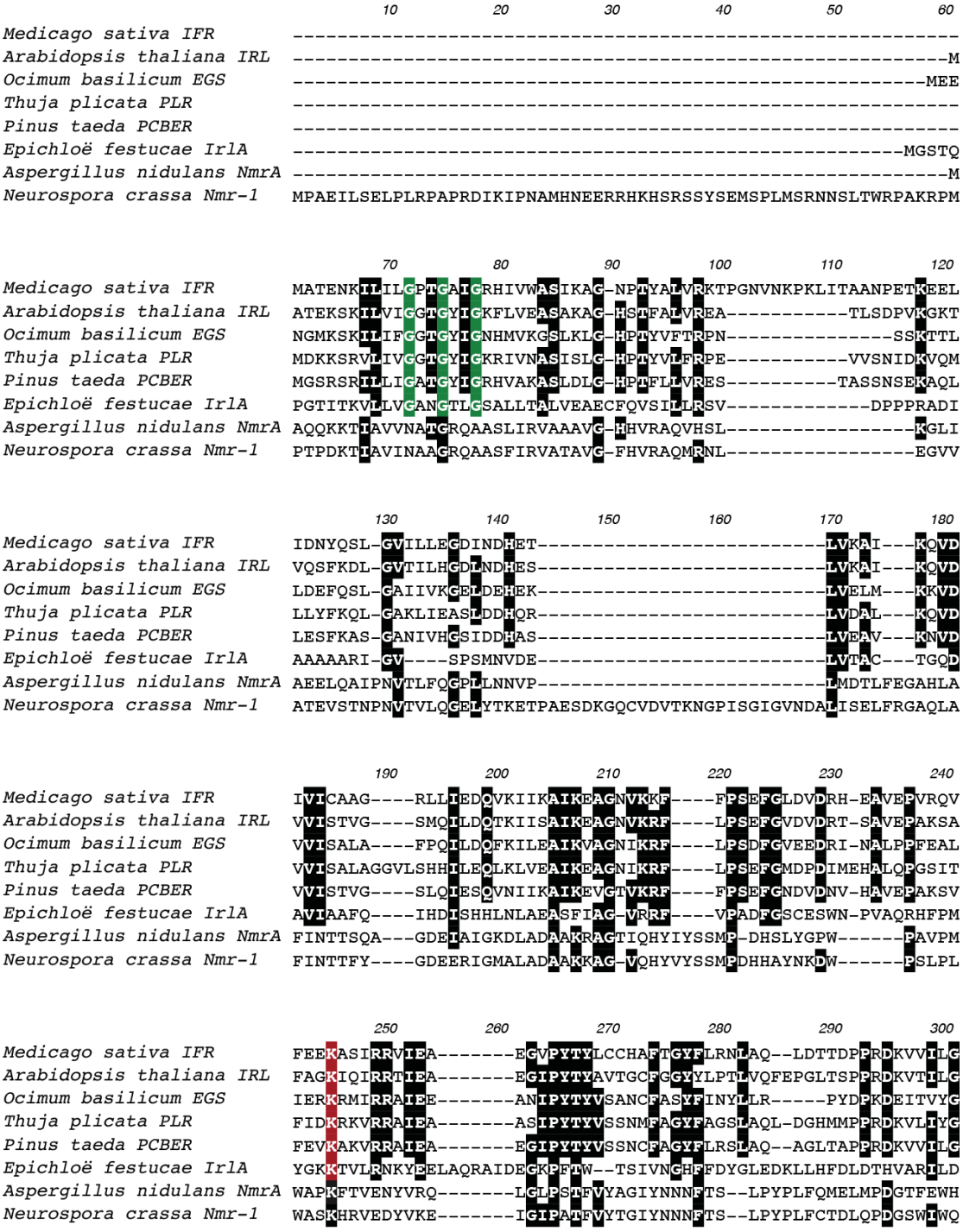
9.4.2 RNA-seq comparison between choked and asymptomatic inflorescences of *E. festucae* E2368-infected *Festuca pratense* plants (tall fescue)[1]

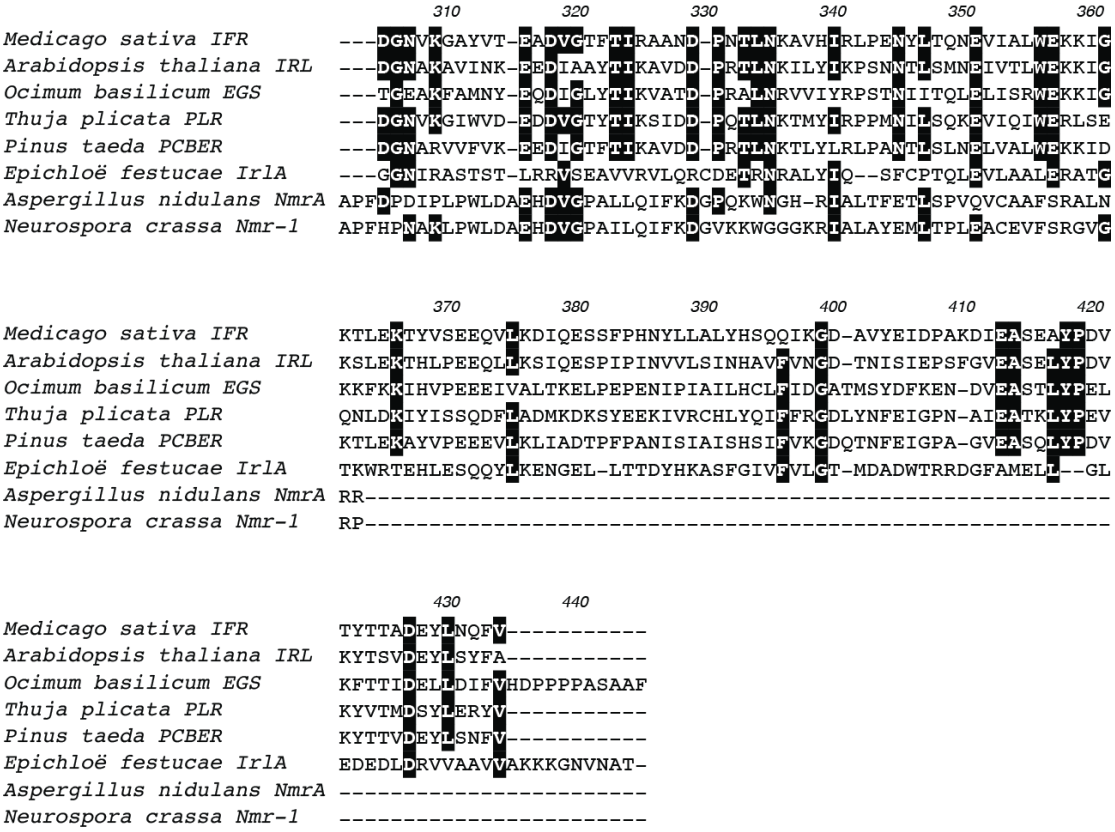
Gene	**Fold change (Stromata / asymptomatic inflorescence; rep 1)**	**Fold change (Stromata / asymptomatic inflorescence; rep 1)**
pdtA	No data	No data
afrA	70	39
irlA	6.1	4.3
mfsB	>300	37
fxbA	>300	240

[1] Data provided by Prof. Chris Schardl, University of Kentucky.

9.5 Multiple sequence alignments

9.5.1 Multiple sequence alignment of proteins related to IrlA[1, 2]





[1] Deposition numbers for sequences: CAA41106.1 (*M. sativa* IFR) and NP_565107.1 (*A. thaliana* IRL), 2QX7_A (*O. basilicum* EGS), Q9LD14.1 (*T. plicata* PLR), 1QYC_A (*P. taeda* PCBER), AAC39442.1 (*A. nidulans* NmrA) and XP_961314.3 (*N. crassa* Nmr-1)

[2] Conserved NADPH binding motif (GXXGXXG) highlighted in green. Conserved IFR/PIP catalytic lysine residue highlighted in red.

9.6 Supplemental nucleotide sequences

9.6.1 Gene sequence for *pdtA* in *E. festucae* Fl1[1]

ATGGCTATTGCGTTTGGAGAGCCGATACCAGTCGGAGAAGAGCATGTAGGTCTCGAAACTTCG
AAACCCCCGGCGTAAGAATCCGGACCTCGGTGCTAATCCGTCCTTACCAGGCCATCTGCGTATCGCTTC
CTGAATGGGACGATTCGATAGGTTTTAGCTGCCAGGATGCAGAGCTCCTGGGTGCCTTGAAAGCAGGAT
ACCCCCGATTCAGCTTGCACCAGGCTGTCGCAAAATTGGAGCAAAACCTGGTGAACTGGTCGACATTTT
CGCCCGGTCCACCGCTGGCCATCGTACCTGGAGAAGCCATCAAGCTGTTCCCGAGTGAGAGGATGGCCA
AGACTTGCCAGCAATACCTGCAGAGAGCTGCAGCAGATGAGAACGGCGGCGAGCGGTGCGTTCGCGTTA
CCCATGTGACCATGGATGGTATACTTGAGGATATCAGTTGTACTCAGTCGCAAGGTCGGCACAAGCACC
CCGAAGAGCACGTCTATGCTGTGACGTACCCTGATCAGCTGGCTCGGCATGCTTGGAAGTTCTGGCTGC
ACACAGGCTACGGCGTTTCCAGCCGATTCTGCGCCTTTTGGCTCAAGAATGCCACCTTTCCCTTCTCTG
AAAATGTAGCGAAACCAAATGGGTTGGCGCGAAAGTTACCCACAACGGAAGCGAAAGAAGCATTCAAAG
CTGTGAGAGAGCATATTGCTAGGCTCTATTCAACTCAAGAGGTCAAGGCGACCGCAGGGGACGTGTTCC
TGTTTCAGGGCGGCATGGCGGCCATTACGCAGACAGCCACGAGCTTAAGCAAGATTGCTGTTGACCGAT
GTGAAAGGCGAAATGGTTACAGGGTCGCTGTATTTGGGTGAGCTCGAGCCTCCTAAATACCGGATAGAC
TCGCCACTGACGTTGTAGTTTCTTATATACTGATACCTTGAAGGTGCTTAGGGACGTCCTTGGCTTTGA
CACTGTCCACTATGGCCACGCATCTTCCGAGGATTTGGATGCACTGGAGCGGGAGTTGAAAACGGGAAG
GTATATTGACGCTCTCTACACCGAGCTCCCCACCAATCCGCTGCTTACGATAGTTGACCTCGAGCGACT
ACATGTTCTGGCGGAGAAATACGACTTTATCCTCGTTGTCGATGATACCATCGGTACCCCCGTAAACCT
CGCAGTAGCTCCCTTGTGCGATGTAGTCTGTACCAGCCTGACCAAGATGTACAGCGGCGCAGGCAACGT
CATGGGCGGAAGTGTCGTCGTAAGTCCCTGGTCCCGTCATCACGATATTTTGTGCGAGACACTTGCAGC
CCACCACGAGGACACCTATTTTCCCCTGGACCTTCTCGTCATGCGCGCCAACAGCGCTCAATTCGAGCA
TCGTGTGACCACGGCAAGCCGCAATGCTCAATGTGTCGCTGAGCTGCTCCGGCGACACAGGCTGGTGGA
AGAGGTCCTCTATCCCGGGGGTGGCCCTACGCAGTCCCTCTACGACAAGGTCAAGCGTCCTGATGCTGG
CTATGGCTTCCTGCTGTCCGTACGCTTCTATCATCCTGAAGCTGCCGTGGCATTCCACGACGCCCTGGA
GACAGCGAAAGGTCCGAGTTTTGGCACCAGCTTCACCCTCGCTTGCCCGTATACCGTGCTCGCACATTA
CGCAGAATTAGAGTGGGCGGCTAGGTTTGGTGTCTGGGAGCACCTAGTCCGTATAAGTGTCGGTATTGA
GCAACCCGAAGTTTTATGCGCGGTTTTTGCAAACGCTCTTGCCGCCGCTGAAAATGCAGCTGCAACTCG
TGGGTAA

[1] CDS shown in black, introns shown in grey

9.6.2 Gene sequence for *afrA* in *E. festucae* Fl1

ATGGCGACTCTGTCATCTACTGCGACTGCGGCGTCAAATCGTGAGCCGAGCAATGCCACGCAT
TGGGATTGCCCATCTCTTGCCGAATCTGCCGCTGTGAAGCATCTGGTGGTGGACGAAGTTGTCAGGCAT
CATCCTGACGGGCAGCTCTTCGACGCACTGGCGATGGCATCGACGTCGGCATCAACCATCATCCATGCA
TCCTGGGCTCTGATTGTCGGTCGTATGACCGATTCTGAGGACGCAATACTTTTTGTCGTCACTGCTGAG
AAGAGCAGTCTGGCGGCACATGTGGCATCAAGCCCAGCACACATCAAGCTGGAGAGCAGCCAGACAATC
GCGCACTATCTTGAAACAGTCCAGAACGAAATCGCAAATTTGACATTTACGGAGCAAGCGATCTCAGAG
ACGACTGAGGGAATACCACTTGGAACTGACAACCAGTCCCCTCAGACGGTGCTTATTGTCAGACATGAA
GATTTGACACCGGCCCAAGACCTTGTCGAGACGTTGTGGACCGAGAGCCAACTACAGAAAATGGCCAGA

CATGCCATTATTCTCGACATACAGATATGTTCGGACCAGTACAAGGCCACAGTGAGTTCCCGTACGGAC
GCCACCACGGCGCCGTGGGCCACCAAGGGGGTGCTTGCACGTCTCGAGCATGTCATCTGCCAGCTTGCA
GGCGCAAGCCGCGATGCAACGTTGCTGGACATTGATATCATCAGTCCACATGACCTGGAGGCTATATGG
ACGTGGAACAGCGACGTGCCTGTGCCAGTCAGAAGATGTGTACATGACTTGATTGCTCAAAAGTCACAA
ATCCAGCCCGATGCGCCAGCCGTCTGCGCTTGGGATGGTGACTTGAGTTATCTCGAACTGGACCAGTAT
GCCACGGCTTTGGCAGTCGAACTGGTCAGCTCCGGTGTCCAAACCGACGTCCTGGTTCCGCTGTGCTTC
GAAAAGTCCAAATGGGCTGTGATAGCCATGCTAGGCGTGCTCAAGGCAGGGGGAGGAGTCGTGTTGCTT
GATCCAAGTCTTCCAGAACAACGACTTGGGTCTATGGTGAAACAAACTGGCGCAGAAATACTGCTTTCC
TCTGCTGCAAATATGGAGTTGAGCACGCGGCTGTGCAAATCGGTTTTCGAAGTTGGTCCAGGTATCACG
AAAATAACAGAGCATGCGTCACCTCCGCACCCGAACACACAGTCACCCTCGACTGCCATGTTTGCTGTA
TTCACATCTGGCAGCACAGGCAGTCCCAAGGGAGTCGTGCTGACACATGAGAACTTCTGCTCTAGCGTG
GAGTATCAGCTCGAGCTTCTCTCATTCACATCCGAGTCTCGCGTCTTTGATTTCGCATCATATGCATTC
GACATCGCCATTCACAACGCCTTTGCAACCTTGGTTGCCGGTGGCTGCCTGTGCGTTCCGTCCGAGAAG
GACCGCAAGGAAAATCTGGCCAAGGTCATGAGGGACATGCGAGCCACCATTGTCGATCTCACGCCGACA
GTGGCTGGTCTTCTAGACCCCGAGACTTTGCCTGATTTGAAGACGATCATTTTTGCCGGAGAGGCCGTC
ACAGTTGAAGACTGTACACCGTGGTGGGGGAAGGCTCAAGTGATCAATGCCTACGGGCCCGCGGAATGC
CACATATGCGTGGTCAACTGCGCCGCCTCGTCCAGCCCTGATGGGGCAACTCGGATAGGCAAGGGAGCT
GGTGTTGTGACTTGGGTTGTCGACCCGGAGAACCATGACCGTCTGCTACCCCCTGGGGGCACGGGCGAG
CTCTTGATTGAAGGGCCGCTTGTTGGACGTGGCTATTTACAGGATGTGAGGAAAACGGAAGCGTCATTC
ATTCACGACCCTGCATGGCTGCTTCGAGGCTCTTCGGCCCACCAAGGACGACAGGGACGGCTTTACAAG
ACGGGCGATCTGGTGCGTTATCACGAGGACGGGAGCCTGTCATACGTCGGTCGCAAGGATACACAAGTC
AAGATCCGTGGACAGCGAGTCGAGCTTGGCGAGGTTGAGATCTGTCTCCAGGAGTGCATGGACAGCAAG
CATCAGGTTGTGGCGGAAGTCTTTTTGCCGCACGAGAACAAGTCGGAACCGTTTTTGGCAGCGTTCCTG
CATTCGAAGAATCATCTTCATACTGGTTCGGAGGCCCAGGTTGCCCTCTTGACCGCTCAATTGAAGGAC
AGATTGGCACAGCGCCTACCAAGCTACATGTGGCCGGTTGTTTTGGTCTACATGAAAGAGTTTCCCGTG
ACAGCTACGGGAAAGATAAACCGCCGAGAATTACGTGAGATGGGACGATTGATGTTTCACAATAGCTCG
CACCTGACCTCTGATGGGTCGAGCAACGGCGAACTGATAGCGGAAACGGAGCAGCCAGCCTATGCGCTG
GCGCAAAAGATTGCCTCCATCGTTCCGGCTTGCCACCGGGCAGGATTAGCAGGACCTGACGGTTTCCGC
GACGTGCTTCTTGCCTCCTCTGGCCTCGACTCAATAAACATGATGTCCCTCATGTACTTCATCTCCAAG
AAGCTTGGTGTGAAGCTCAGCATGAAGCAGCTCGCGGACGAGACCATGACTATTCGCGCTTTGGCGAGG
TGTGTGGCGGAGAAGCAAACAGCTGCGGTCTCTCATGAGCCCAAAAACTCGGTACAAACCGGCGGCATC
GACACTCTGGCCGAGATCAACCGTCTAGACTCTCGTATCTCTGAACATCAGAAGCTGAATTGTGCCATG
ACAGACGTCGACTCACCAAAGAAGACTGGTTCCAATCTGATCGTCTTTCTCACAGGCGCCAATGGCTAT
CTGGGCACTCAGATCTTACGACAGCTCCTCGAGCACAGAAATGTAGACCGCGTTATTGCCTTGGTACGC
GGCACATCGCCAGCCTCGGCAAGAACGCGTGTCATTGATGCGGCAAAAAAATCACTTTGGTGGACAGAC
GTGCACGGCGAGAAGCTTGATGTCTGGGTAGGAGACCTCTCATTGCCGTGCTTGGGGCTCGATTCTACC
AACTGGGCGCTTCTCAAGGATGGCAAGTCAGTTGATGTCATCATTCACAACGGCGCAACGGTGCACTGG
GCAAAGAGCTACTCGGCACTCGAGGCCACCAATGTCCGCTCAACCGTCCATCTGCTGCTCATTGCACTG
TCATCGCCTCGGGCGAAGTTTGTCTACGTGTCCAACCGGCGCACAGCCGGCTCGGAAAACCTGACAGAG
GCGGACATAGCCCGCGGTCTGGTACACGTGGACTCGGTCGCGTACAGCCAGACCAAATTCGTTGCGGAG
GCGTTGGTCAAGCGAGCTGCGGACCGTGACGTCTCGGCAGCACGGCGCCTCGCCGTGGTGAGCCCCGGC
TACGTTGTAGGCACGCCTGCCGAGGGCATTGGCAACGCGGATGACTACGTCTGGCGGCTTGTGGCAACT

TGCGTCAGGGTGGGCGCGTACAACATTGACGAGGGGGACACATGGCTGCCTCTGGCTGACGCAACGACC
GCGGCGGCCGTCATCACCGGAGCAGCGTTCCGATGGGAATCCGAGGCCAAGACGGTGACGCCAGTCATG
GACGGTATAACCTGGGCTGGAATCTGGAGCATCATGAGCGGCATGGGATATCGGCTCCGAGGCATGAAC
GCGAACGAGTGGCTGGCCGCTGCTCTGGCCGACGTTCAACGAGTCGGTGAGAGACACCCCTTTTGGCCG
CTGGCGCACCTGCTCGACGGTCAAAAGCTTCAGGGGAAGAAGACGGGGGGAACATGGGATCAGACTGGT
CACACCCCGTTGGTATTGAAGGTTGCCCTGAGAAAAACAGCAGAGTCTTTGTGCAGGGCCGGGTTTCTG
CCCCTGCCCCTGCCCCTGCCTAATCACGGAGAATGCGGTGACACATTGGATAGGCACGAAGAAAAGCCT
CGGCTTGAGTGTGATGTAGTTTTTCGGCGAGCTGCTCCCGGCATCTCGGAGCAAGTTGCGGTGGCCACG
ACCGTGTAA

9.6.3 Gene sequence for *irlA* in *E. festucae* Fl1

ATGGGGTCAACCCAACCAGGCACAATCACCAAAGTCCTCCTTGTCGGAGCCAATGGCACTCTG
GGCAGTGCTCTTCTCACGGCCCTCGTGGAAGCAGAGTGCTTCCAAGTTTCCATCTTGCTGCGTTCGGTA
GATCCTCCGCCTCGGGCAGACATCGCGGCCGCGGCGGCGAGAATAGGTGTCTCGCCGAGCATGAATGTC
GATGAACTGGTCACGGCCTGCACTGGCCAGGACGCCGTCATCGCCGCATTCCAAATCCACGACATCTCC
CATCACCTGAACCTTGCAGAGGCCTCCTTCATCGCGGGAGTCCGGCGCTTCGTTCCCGCCGACTTTGGC
AGCTGTGAGTCTTGGAACCCGGTCGCGCAGCGCCATTTCCCAATGTACGGCAAAAAGACTGTCCTGCGC
AACAAGTACGAGGAGCTTGCTCAAAGAGCCATCGATGAAGGCAAGCCGTTCACCTGGACTTCGATTGTC
AACGGACACTTTTTCGACTACGGACTCGAGGACAAACTCTTGCACTTCGACCTGGACACCCACGTAGCG
CGGATTCTCGACGGGGGCAACATCAGAGCGTCAACCTCGACACTGCGGCGGGTTTCCGAGGCAGTCGTG
CGGGTGCTGCAGAGGTGTGACGAGACTCGGAACCGAGCTTTGTACATCCAGAGCTTCTGTCCCACGCAG
CTCGAGGTGCTGGCTGCGTTGGAGCGTGCCACGGGCACCAAGTGGCGTACTGAGCATCTCGAGTCCCAG
CAGTATCTCAAGGAGAATGGTGAACTGCTCACTACGGACTATCACAAGGCCAGTTTTGGTATTGTGTTT
GTGCTGGGAACCATGGACGCAGACTGGACACGACGGGACGGTTTTGCTATGGAATTATTGGGACTGGAG
GACGAGGACTTGGATCGTGTCGTGGCTGCGGTTGTGGCCAAGAAGAAGGGGAACGTGAATGCTACATAG

9.6.4 Gene sequence for *mfsB* in *E. festucae* Fl1[1]

ATGACGGCAAACGGCAAACCATGGGGTCTTCATTGGCGCTCGATGCCATCCTTCATCGTAACG
ACGGTGGCGATCGGGCTCTTCACAGATCTCTTTCTGTACGGCATCATGGTGCCGATATTGCCCTTTCTC
CTGCAGGACCGTCTTTCTATTCCTGGCAATCAGATCCAGTCGTATACCTCCGCTTTACTTGCGGCGTAT
GCTGGGCCATCACTCCTGTTTGCTATTCCAGCTGGATGGGTTACGGACAAGATTGGCGCCCGTCAGCGG
CCTTTCCTCGCCGGCTTGCTGTTGCAGTTGTTCGCCACAACCGCGATGACGTTTGGAAAGACCATGACG
GTGCTGGTCATCGCAAGGCTGGTCCAGGGGTTGGCTGCCCCGGTGGTTTGGACGACGGGCCTTGCCATG
GTCCAAGAAACAACTGACCCTGACAGAATCGGCGAGGCCATTGGCGTGGTAAGCATCTTGTACTTTGTT
CCTTACAAGAGAGAGATCTGACAAAGTTTCATAATAGATTTTTGCGGTAATCTCTGTTGGCGAAGTCGC
GGCCCCGGTCGTCGGAGGCGTACTCTACGACGAGGTTGGCATCGCCGGTGTTTTCGGCGTGGCAATAGG
AATTCTAAGTGTTGACGTCGTCATGCGAATTCTCGTGCTCGACAAGAAGACGTTGGCCGGGTACGAGGC
GGCAAACATGCACGATCCCGCAAGTAATGCTGCTGCAGAGGAACGCGAGCGCGAACCCCTGCTTGCTTC
TCCTCCTCCTTCCAATGATGCATACAAAATCCGCGAACAGGTCGGCAGTTTTACCCGAGCCGTCCCCAT
CTTCTACTGCTTCCGTGAGCCCCGGTTCATCATGGCCTTCATGATGACGCTGGTCCAGGCCACGTTCAG
AGGAATATTCGACGCCACGGTTCCAATCGAGGCACAATCGTTATTCCACTTCTCCTCTCAGACGGTCGG
CCTTTTGTTCATCACGCTCTTTGTGCCGCATTTGACACTGAGCCATGTCGTCGGCCGGTCGGTGGACAA

GTACGGAACAAGAGCAGTGGCCACGGCCGGATTCGTCTTGTTAGTCCCCTGTCTTGCCGCGCTGGGTCT
TCCTTCGCAACAGTGCGTCGCCGGAAGCGCAAAGGTCGTCTTGTTCTGCGTCATACTGGCCTTGAACGG
TATTGGGTTGTGCATGGTCAGTACACCGGCCTTCGTGGAAGCTATACGGGTCACGCAGAGGTACGAGGC
ATCCAATCCGGGATTCTTTGGGGAATACGGGCCCTATGCTCAACTGTACGGGTTCAATTGCCTCTTCTT
CTCTGCTGGGCTTACAATCGGGCCGATAGTGGGCGGCGTTCTGAGAGATGTTTTTGGGTACGGAGTTGT
CGGCATTCTGTTTTCTGTCGTCTCTGGCGTTACAGCCGCATTGGCTTTTACAATCATTGGCGATGGACG
GCGGTGA

[1] CDS shown in black, introns shown in grey

9.6.5 Gene sequence for *fxbA* in *E. festucae* Fl1[1]

ATGACAACTTGGGTGAGCAAAATCGATCTGGATTCAGATCAATCTGCTCTACTCGATGCGCAA
AGACCCGGAGAGGCTAGGACGCGAATAAAAACGTGCCGTAGTTTCGGCAAGCGGTTATCCGCTGCGGGA
GAATGTAAAGGTTCAAAAAGCGGAACGAGGGCGACAAATGTGTCACTGTGGCCGACAGTTGCTCTTCTT
TTATGCGTCCTATCCGCCTTCCGTCTGGCGAGGAAGCATCTCACCGATCGAGACGAACTGGCCGCTTGC
CTTTCTTCCAAGCAGATTCCGTTTGTAGTCGAAAGCTCAGATGCGTGGAGTCAAAAGGCTGTGACTTGG
AATATTCGACTGGCATACAAGCCCGCCGCAATCGCAATACCGCGATCTACTCGTCATGTCGCAGCTGCG
GTAACATGCGGTGTCAACAACGGCGTGCGAGTGAGTGCCAAGAGTGGTGGCCACAGCTTTGGATCCTAT
GGCTTTGGGGGCGAGGATGGGCATCTGGTGATTGACCTTGGTCTGATGGACGGGATTGTCCTCGACGAC
GATGGATCTGCGAAGATTCAACCAGGTGCCCGGCTAGGGCGTGTGGCGATGGAACTGTACAATCAAGGC
GGCAGGGCTATCCCGCATGGTTCATGCCCGGGGTAAGTGTCCCGCTCCTTTAAATTGATGATTCAAACT
TGATGAAACTTGACTCTTGTACTAGATAGCTGATGGATGCAGGGTCGGCGTTGCCGGGCATGTTTTGCA
CGGCGGCTACGGGATTGCCTCGCGACTTCACGGCCTCACATTGGACTGGTTGATCGGTGCAACCGTTGT
TCTCGCCAATGGTTCGAGAGCGCACTGCTCTGAGACGGACAACACAGACCTGTTCTGGGCATTGCGTGG
TGCAGGGTCATCGTTTGGCATTGTTACCGAGTTCGAGTTCAAGACTTTCGAAGCACCCGACCAGGTCAC
CCCGTTTACCATCGATCTACACTGGAACGAGGCGCAAGCAGTTCAGGGATTGAAAACTATTCAAGAGTT
CGCCATGACAGCTGCAAGAGAGCTCAACTTGTGGCTCTTTATGGGCAGAGTGAGCCAAAGCATTCAAGG
CGTGTACTTTGGGCACGTCACTGGTCTTAACGAGGCGATCCAGCCCTTTCTCGACCGTACTGGAGCGGA
AGTTTCGTATGCAACCACTATGGGATGGATACAGGCGCATGCGCATTTCGCAAATGGTACCGACCTCAG
ACAAACGTCTGCATATGACTTGGTATGCCGCCTTGGTTCTATTGAAACGACGGCAGATATGGCGGAACT
AACTAAATCTCAGCACAAGAGGTTCTATGGGACTAGTCTCATGGCCAATGCCCTCACAGATGATCACCT
CAAGTCGTTCATGTCGGCAATCTTTAGCAACATCGACGAAGTGTCAGCGCGCCATTCCTGGGAACTCAC
AGTTGCGTTTCACGGAGGTGATGGCTCAGCAATCGCAGATGTCAATACATCTGCAACCGCATACGTCCA
TCGCGACAAGCTGCTTCTTTATCAGTTTTCTGACATTGAAACTCACGGCCACTATCCCGAGGATGGTTT
CGCTGTCCTACGGCGGTTCAGAGAGAGTGTTACCAATTCGTTGGAGGATAGACAGTGGGGGATGTACGC
GAACTATGTGGACACACAGTTGGATTCAAAGACGGCACAGAAGCTGTATTGGGGCAATAATCTCCATCG
ATTACAGCATACCAAATCAGCTTTGGATCCCGATGAGGTGTTTTGGGACCCGCACAGCATTCGTCCTCT
GCCCTAA

[1] CDS shown in black, introns shown in grey

9.6.6 Gene sequence for *tesB* in *E. festucae* Fl1[1]

ATGGCTCGCGAGGAACGGTCCACGCTTTTGCGGCCACCGCCAGCGGATCCCGACAAGGCGCCGATC
GAGAACGTGCTGGAGGTGACGGAAGTAGGGGTTCTAGGCACGGTAAGTTTGTGTTTCGAGCGAGAACAACAG
ATGGGATCCAGATTTTGTCGAGAAAACAAGCACAGAAAAAAAAAAAACTCACGCCGGCAAAAAAAAAAAAAAA
AGGACATCTTCACAAACACGCGCAAGGCATGGGTTCCCCCCGGCGCAAGAGGCGTGTACGGCGGCACCGTCAT
CGCGCAGTGTCTCGCGTCGGCGCAGAAGACTGTCCCCGACGACTTCCACCCGCACTCGTGCCACTGCTACTTCA
TCCTCGCCGGGTCGGGGGCCGTCCCGATCCTGTACCACGTCGAGCGCGTGCGGGACGGGCGCTCGTTCGCCAC
GAGGACGGTGCAGGCGCGGCAGCGGGGCCGCTGCATCTTCACCACGACCATCAGCTTCGTGCGCGAGGCCGA
GGCGGAGGCGGAGGCGGCGGCGGGCGGGGAGAGGAGGGGGACGGGGACGGGGAGGACGGTGGTGGTGG
TTCCCACGGTCCGGCACGCGACGCCGATGCCCGAGGAGGCGCACTCGCCCCCGCCCGAGGACTGGGACGACG
AGCCCGGGTGGGCGAGGGGCGGCCCGTTTGTCAGCTACAGGATGCCCGACGTGCGGAGGAGGGGGCACGGC
GCGGACGGCGCGGACGAGCGGCCGGACCAGGTGCGCTGCCAGCAGTGGGTGAGGTGCCGAGGGAGCATCTC
CCGCGCGGGCGGGCGGCAGGCTCACCTCAACGCGCTGGCGTACGTCTCGGACAGCGCCTTCATCAGCACCGTG
TCGCGCGTCCACAGGCTGTGGCGGCTGCCGTTCGCGCCCGACGACTACGAGTCCATGGACGACGAGCAACGG
GCTCGCGTCAGGGACTTTGCGCGGCAAGACGGCCTCGGGACCGCGGTCGAGGAGTGGAAGCAGAGGGGGAG
GCTGGGCATGATGGTCAGCCTGGACCACAGCATATACTTCCACGAGCCGGCGAGCGTCAGGGTCGACGAATG
GATGTTTGCCGACATGGACAGTCCGTGGTCGGGCGATGGGAGGGGGTTCGTCACGCAGCGCATCTTTGCGCGC
GACGGCACGTTGCTGGCGACGTGTGTGCAGGAGGTGAGCTTCCTTTCGCCCTGTTTTTTAACGCGTCCTTCATG
TCCTTCATGTCCTTCATGGCAGCTTTCATCTCTACTAGTTCAGTGCTAATCACCCCCCCAGGGCGTCATCCGACTC
AGACCATCTGAAAAGGAAAAGACAAAGTTGTAA

[1] CDS shown in black, introns shown in grey

10 References

Ambrose, K.V., Koppenhöfer, A.M., and Belanger, F.C. (2014). Horizontal gene transfer of a bacterial insect toxin gene into the *Epichloë* fungal symbionts of grasses. Scientific Reports *4*, 5562.

Andrianopoulos, A., Kourambas, S., Sharp, J.A., Davis, M.A., and Hynes, M.J. (1998). Characterization of the *Aspergillus nidulans nmrA* gene involved in nitrogen metabolite repression. J Bacteriol *180*, 1973-1977.

Ansari, M.Z., Sharma, J., Gokhale, R.S., and Mohanty, D. (2008). In silico analysis of methyltransferase domains involved in biosynthesis of secondary metabolites. BMC Bioinformatics *9*.

Arachevaleta, M., Bacon, C., Hoveland, C., and Radcliffe, D. (1989). Effect of the tall fescue endophyte on plant response to environmental stress. Agron J *81*, 83-90.

Aral, B., and Kamoun, P. (1997). The proline biosynthesis in living organisms. Amino Acids *13*, 189-217.

Bacon, C.W. (1988). Procedure for isolating the endophyte from tall fescue and screening isolates for ergot alkaloids. Appl Environ Microbiol *54*, 2615-2618.

Balibar, C.J., Howard-Jones, A.R., and Walsh, C.T. (2007). Terrequinone A biosynthesis through L-tryptophan oxidation, dimerization and bisprenylation. Nat Chem Biol *3*, 584-592.

Ball, O.J.P., Miles, C.O., and Prestidge, R.A. (1997). Ergopeptine alkaloids and *Neotyphodium lolii*-mediated resistance in perennial ryegrass against adult *Heteronychus arator* (Coleoptera: Scarabaeidae). J Econ Entomol *90*, 1382-1391.

Becker, M., Becker, Y., Green, K., and Scott, B. (2016). The endophytic symbiont *Epichloë festucae* establishes an epiphyllous net on the surface of *Lolium perenne* leaves by development of an expressorium, an appressorium-like leaf exit structure. New Phytol.

Becker, Y., Eaton, C.J., Brasell, E., May, K.J., Becker, M., Hassing, B., Cartwright, G.M., Reinhold, L., and Scott, B. (2015). The fungal cell-wall integrity MAPK cascade is crucial for hyphal network formation and maintenance of restrictive growth of *Epichloë festucae* in symbiosis with *Lolium perenne*. Mol Plant-Microbe Interact *28*, 69-85.

Berry, D., Cox, M.P., and Scott, B. (2015a). Draft genome sequence of the filamentous fungus *Penicillium paxilli* (ATCC 26601). Genome Announcements *3*, e00071-00015.

Berry, D., Takach, J.E., Schardl, C.L., Charlton, N.D., Scott, B., and Young, C.A. (2015b). Disparate independent genetic events disrupt the secondary metabolism gene *perA* in certain symbiotic *Epichloë* species. Appl Environ Microbiol *21*, 2798-2807.

Bhattacharjee, J.K. (1985). α-Aminoadipate pathway for the biosynthesis of lysine in lower eukaryotes. Crit Rev Microbiol *12*, 131-151.

Brem, D., and Leuchtmann, A. (1999). High prevalence of horizontal transmission of the fungal endophyte *Epichloë sylvatica*. Bull Geobot Inst ETH, 3-12.

Bultman, T.L., and Leuchtmann, A. (2008). Biology of the *Epichloë–Botanophila* interaction: an intriguing association between fungi and insects. Fungal Biology Reviews *22*, 131-138.

Bushley, K.E., and Turgeon, B. (2010). Phylogenomics reveals subfamilies of fungal nonribosomal peptide synthetases and their evolutionary relationships. BMC Evol Biol *10*, 1-23.

Byrd, A.D., Schardl, C.L., Songlin, P.J., Mogen, K.L., and Siegel, M.R. (1990). The β-tubulin gene of *Epichloë typhina* from perennial ryegrass (*Lolium perenne*). Curr Genet *18*, 347-354.

Caboche, S., Leclere, V., Pupin, M., Kucherov, G., and Jacques, P. (2010). Diversity of monomers in nonribosomal peptides: Towards the prediction of origin and biological activity. J Bacteriol *192*, 5143-5150.

Caboche, S., Pupin, M., Leclere, V., Fontaine, A., Jacques, P., and Kucherov, G. (2008). NORINE: a database of nonribosomal peptides. Nucleic Acids Res *36*, D326-D331.

Casqueiro, J., Gutierrez, S., Banuelos, O., Fierro, F., Velasco, J., and Martin, J.F. (1998). Characterization of the *lys2* gene of *Penicillium chrysogenum* encoding α-aminoadipic acid reductase. Mol Gen Genet *259*, 549-556.

Challis, G.L., Ravel, J., and Townsend, C.A. (2000). Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Chem Biol *7*, 211-224.

Charlton, N.D., Shoji, J.-Y., Ghimire, S.R., Nakashima, J., and Craven, K.D. (2012). Deletion of the fungal gene soft disrupts mutualistic symbiosis between the grass endophyte *Epichloë festucae* and the host plant. Eukaryot Cell *11*, 1463-1471.

Chiang, Y.-M., Szewczyk, E., Davidson, A.D., Keller, N., Oakley, B.R., and Wang, C.C.C. (2009). A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*. J Am Chem Soc *131*, 2965-2970.

Christensen, M.J., Bennett, R.J., Ansari, H.A., Koga, H., Johnson, R.D., Bryan, G.T., Simpson, W.R., Koolaard, J.P., Nickless, E.M., and Voisey, C.R. (2008). *Epichloë* endophytes grow by intercalary hyphal extension in elongating grass leaves. Fungal Genet Biol *45*, 84-93.

Christensen, M.J., Leuchtmann, A., Rowan, D.D., and Tapper, B.A. (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.

Christensen, M.J., and Voisey, C.R. (2007). The biology of the endophyte/grass partnership. Paper presented at: the 6th International Symposium on Fungal Endophytes of Grasses (Christchurch, New Zealand: New Zealand Grassland Association).

Chujo, T., and Scott, B. (2014). Histone H3K9 and H3K27 methylation regulates fungal alkaloid biosynthesis in a fungal endophyte-plant symbiosis. Mol Microbiol *92*, 413-434.

Chung, K.R., and Schardl, C.L. (1997). Sexual cycle and horizontal transmission of the grass symbiont, *Epichloë typhina*. Mycol Res *101*, 295-301.

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

Conti, E., Stachelhaus, T., Marahiel, M.A., and Brick, P. (1997). Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. EMBO J *16*, 4174-4183.

Cox, R.J., and Simpson, T.J. (2009). Chapter 3: fungal type I polyketide synthases. In Methods Enzymol (Academic Press), pp. 49-78.

di Salvo, M.L., Budisa, N., and Contestabile, R. (2013). PLP-dependent enzymes: a powerful tool for metabolic synthesis of non-canonical amino acids. Paper presented at: Beilstein Bozen symposium on molecular engineering and control (Beilstein Institute, Prien, Germany).

Donzelli, B.G.G., Krasnoff, S.B., Sun-Moon, Y., Churchill, A.C.L., and Gibson, D.M. (2012). Genetic basis of destruxin production in the entomopathogen *Metarhizium robertsii*. Curr Genet *58*, 105-116.

Drozdetskiy, A., Cole, C., Procter, J., and Barton, G.J. (2015). JPred4: a protein secondary structure prediction server. Nucleic Acids Res *43*, W389-W394.

Du, L., and Lou, L. (2010). PKS and NRPS release mechanisms. Nat Prod Rep *27*, 255-278.

Dupont, P.-Y., Eaton, C.J., Wargent, J.J., Fechtner, S., Solomon, P., Schmid, J., Day, R.C., Scott, B., and Cox, M.P. (2015). Fungal endophyte infection of ryegrass reprograms host metabolism and alters development. New Phytol *208*, 1227-1240.

Eaton, C.J., Cox, M.P., Ambrose, B., Becker, M., Hesse, U., Schardl, C.L., and Scott, B. (2010). Disruption of signaling in a fungal-grass symbiosis leads to pathogenesis. Plant Physiol *153*, 1780-1794.

Eaton, C.J., Dupont, P.-Y., Solomon, P., Clayton, W., Scott, B., and Cox, M.P. (2015). A core gene set describes the molecular basis of mutualism and antagonism in *Epichloë* spp. Mol Plant-Microbe Interact *28*, 218-231.

Eaton, C.J., Jourdain, I., Foster, S.J., Hyams, J.S., and Scott, B. (2008). Functional analysis of a fungal endophyte stress-activated MAP kinase. Curr Genet *53*, 163-174.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res *32*, 1792-1797.

Ehmann, D.E., Trauger, J.W., Stachelhaus, T., and Walsh, C.T. (2000). Aminoacyl-SNACs as small-molecule substrates for the condensation domains of nonribosomal peptide synthetases. Chem Biol *7*, 765-772.

Eisfeld, K. (2009). Non-ribosomal peptide synthetases of fungi. In Physiology and Genetics, T. Anke, and D. Weber, eds. (Springer Berlin Heidelberg), pp. 305-330.

Ewald, P.W. (1987). Transmission modes and evolution of the parasitism-mutualism continuum. Ann N Y Acad Sci *503*, 295-306.

Finking, R., and Marahiel, M.A. (2004). Biosynthesis of nonribosomal peptides. Annu Rev Microbiol *58*, 453-488.

Fleetwood, D.J., Khan, A.K., Johnson, R.D., Young, C.A., Mittal, S., Wrenn, R.E., Hesse, U., Foster, S.J., Schardl, C.L., and Scott, B. (2011). Abundant degenerate miniature inverted-repeat transposable elements in genomes of epichloid fungal endophytes of grasses. Genome Biol Evol *3*, 1253-1264.

Fleetwood, D.J., Scott, B., Lane, G.A., Tanaka, A., and Johnson, R.D. (2007). A complex ergovaline gene cluster in *Epichloë* endophytes of grasses. Appl Environ Microbiol *73*, 2571-2579.

Forseth, R.R., Amaike, S., Schwenk, D., Affeldt, K.J., Hoffmeister, D., Schroeder, F.C., and Keller, N.P. (2013). Homologous nrps-like gene clusters mediate redundant small-molecule biosynthesis in *Aspergillus flavus*. Angew Chem *52*, 1590-1594.

Fu, Y.H., Young, J.L., and Marzluf, G.A. (1988). Molecular cloning and characterization of a negative-acting nitrogen regulatory gene of *Neurospora crassa*. Molecular & General Genetics *214*, 74-79.

Gacek, A., and Strauss, J. (2012). The chromatin code of fungal secondary metabolite gene clusters. Appl Microbiol Biotechnol *95*, 1389-1404.

Gardiner, D.M., and Howlett, B.J. (2004). Negative selection using thymidine kinase increases the efficiency of recovery of transformants with targeted genes in the filamentous fungus *Leptosphaeria maculans*. Curr Genet *45*, 249-255.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., III, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods *6*, 343-345.

Glinski, M., Hornbogen, T., and Zocher, R. (2001). Enzymatic synthesis of fungal N-methylated cyclopeptides and depsipeptides. In Enzyme technologies for pharmaceutical and biotechnological applications, H.A. Kirst, Wu-Kuang Yeh, ed. (Dekker, New York: CRC Press), pp. 471-497.

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

Hacker, C., Glinski, M., Hornbogen, T., Doller, A., and Zocher, R. (2000). Mutational analysis of the N-methyltransferase domain of the multifunctional enzyme enniatin synthetase. J Biol Chem *275*, 30826-30832.

Hahn, H., McManus, M.T., Warnstorff, K., Monahan, B.J., Young, C.A., Davies, E., Tapper, B.A., and Scott, B. (2008). *Neotyphodium* fungal endophytes confer physiological protection to perennial ryegrass (*Lolium perenne* L.) subjected to a water deficit. Environ Exp Bot *63*, 183-199.

Hahn, M., and Stachelhaus, T. (2004). Selective interaction between nonribosomal peptide synthetases is facilitated by short communication-mediating domains. Proc Natl Acad Sci U S A *101*, 15585-15590.

Haslinger, K., Redfield, C., and Cryle, M.J. (2015). Structure of the terminal PCP domain of the non-ribosomal peptide synthetase in teicoplanin biosynthesis. Proteins: Struct Funct Bioinform *83*, 711-721.

Hua, C., Linling, L., Feng, X., Yan, W., Honghui, Y., Conghua, W., Shaobing, W., Zhiqin, L., Juan, H., Yuping, W*., et al.* (2013). Expression patterns of an isoflavone reductase-like gene and its possible roles in secondary metabolism in Ginkgo biloba. Plant Cell Rep *32*, 637-650.

Hur, G.H., Vickery, C.R., and Burkart, M.D. (2012). Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology. Nat Prod Rep *29*, 1074-1098.

Itoh, Y., Johnson, R., and Scott, B. (1994). Integrative transformation of the mycotoxin-producing fungus, *Penicillium paxilli*. Curr Genet *25*, 508-513.

Ji, Y.-l., Zhan, L.-h., Kang, Y., Sun, X.-h., Yu, H.-s., and Wang, Z.-w. (2009). A new stromata-producing *Neotyphodium* species symbiotic with clonal grass *Calamagrostis epigeios* (L.) Roth. grown in China. Mycologia *101*, 200-205.

Johnson, L.J., de Bonth, A.C.M., Briggs, L.R., Caradus, J.R., Finch, S.C., Fleetwood, D.J., Fletcher, L.R., Hume, D.E., Johnson, R.D., Popay, A.J*., et al.* (2013a). The exploitation of epichloae endophytes for agricultural benefit. Fungal Diversity *60*, 171-188.

Johnson, L.J., Johnson, R.D., Schardl, C.L., and Panaccione, D.G. (2003). Identification of differentially expressed genes in the mutualistic association of tall fescue with *Neotyphodium coenophialum*. Physiol Mol Plant Pathol *63*, 305-317.

Johnson, L.J., Koulman, A., Christensen, M., Lane, G.A., Fraser, K., Forester, N., Johnson, R.D., Bryan, G.T., and Rasmussen, S. (2013b). An extracellular siderophore is required to maintain the mutualistic interaction of *Epichloë festucae* with *Lolium perenne*. PLoS Path *9*, e1003332.

Johnson, R., Borchert, S., Christensen, M., Johnson, L., Koulman, A., van Gils, M., and Bryan, G. (2007a). A gene identified from *Neotyphodium lolii* is expressed only *in planta* and regulates the biosynthesis of a putative oligopeptide secondary metabolite. Paper presented at: Proceedings of the 6th International Symposium on Fungal Endophytes of Grasses (Christchurch, New Zeland).

Johnson, R., Voisey, C., Johnson, L., Pratt, J., Fleetwood, D., Khan, A., and Bryan, G. (2007b). Distribution of NRPS gene families within the *Neotyphodium*/*Epichloë* complex. Fungal Genet Biol *44*, 1180-1190.

Johnson, R.D., Lane, G.A., Koulman, A., Cao, M., Fraser, K., Fleetwood, D.J., Voisey, C.R., Dyer, J.M., Pratt, J., Christensen, M*., et al.* (2015). A novel family of cyclic oligopeptides derived from ribosomal peptide synthesis of an in planta-induced gene, *gigA*, in *Epichloë* endophytes of grasses. Fungal Genet Biol *85*, 14-24.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G.*, et al.* (2014). InterProScan 5: genome-scale protein function classification. Bioinformatics *30*, 1236-1240.

Kalb, D., Lackner, G., and Hoffmeister, D. (2013). Fungal peptide synthetases: an update on functions and specificity signatures. Fungal Biology Reviews *27*, 43-50.

Kalb, D., Lackner, G., and Hoffmeister, D. (2014). Functional and phylogenetic divergence of fungal adenylate-forming reductases. Appl Environ Microbiol *80*, 6175-6183.

Kalb, D., Lackner, G., Rappe, M., and Hoffmeister, D. (2015). Activity of α-aminoadipate reductase depends on the N-terminally extending domain. ChemBioChem *16*, 1426-1430.

Kayano, Y., Tanaka, A., Akano, F., Scott, B., and Takemoto, D. (2013). Differential roles of NADPH oxidases and associated regulators in polarized growth, conidiation and hyphal fusion in the symbiotic fungus *Epichloë festucae*. Fungal Genet Biol *56*, 87-97.

Keating, T.A., Marshall, C.G., Walsh, C.T., and Keating, A.E. (2002). The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains. Nat Struct Biol *9*, 522-526.

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

Keller, N.P., Turner, G., and Bennett, J.W. (2005). Fungal secondary metabolism - from biochemistry to genomics. Nat Rev Microbiol *3*, 937-947.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protocols *10*, 845-858.

Koeduka, T., Fridman, E., Gang, D.R., Vassão, D.G., Jackson, B.L., Kish, C.M., Orlova, I., Spassova, S.M., Lewis, N.G., and Noel, J.P. (2006). Eugenol and isoeugenol, characteristic aromatic constituents of spices, are biosynthesized via reduction of a coniferyl alcohol ester. Proc Natl Acad Sci USA *103*, 10128-10133.

Koglin, A., Loehr, F., Bernhard, F., Rogov, V.V., Frueh, D.P., Strieter, E.R., Mofid, M.R., Guentert, P., Wagner, G., Walsh, C.T.*, et al.* (2008). Structural basis for the selectivity of the external thioesterase of the surfactin synthetase. Nature *454*, 907-911.

Koglin, A., Mofid, M.R., Lohr, F., Schafer, B., Rogov, V.V., Blum, M.M., Mittag, T., Marahiel, M.A., Bernhard, F., and Dotsch, V. (2006). Conformational switches modulate protein interactions in peptide antibiotic synthetases. Science *312*, 273-276.

Kotowska, M., and Pawlik, K. (2014). Roles of type II thioesterases and their application for secondary metabolite yield improvement. Appl Microbiol Biotechnol *98*, 7735-7746.

Krappmann, S., Sasse, C., and Braus, G.H. (2006). Gene targeting in *Aspergillus fumigatus* by homologous recombination is facilitated in a nonhomologous end-joining-deficient genetic background. Eukaryot Cell *5*, 212-215.

Lambalot, R.H., Gehring, A.M., Flugel, R.S., Zuber, P., LaCelle, M., Marahiel, M.A., Reid, R., Khosla, C., and Walsh, C.T. (1996). A new enzyme superfamily - The phosphopantetheinyl transferases. Chem Biol *3*, 923-936.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R.*, et al.* (2007). Clustal W and clustal X version 2.0. Bioinformatics *23*, 2947-2948.

Latch, G.C.M., and Christensen, M.J. (1985). Artificial infection of grasses with endophytes. Ann Appl Biol *107*, 17-24.

Lautru, S., and Challis, G.L. (2004). Substrate recognition by nonribosomal peptide synthetase multi-enzymes. Microbiology-Sgm *150*, 1629-1636.

Lee, T.V. (2008). Structural biology of endophyte non-ribosomal peptide synthetases (Auckland, New Zealand: University of Auckland).

Lee, T.V., Johnson, L.J., Johnson, R.D., Koulman, A., Lane, G.A., Lott, J.S., and Arcus, V.L. (2010). Structure of a eukaryotic nonribosomal peptide synthetase adenylation domain that activates a large hydroxamate amino acid in siderophore biosynthesis. J Biol Chem *285*, 2415-2427.

Leuchtmann, A., Bacon, C.W., Schardl, C.L., White, J.F., Jr., and Tadych, M. (2014). Nomenclatural realignment of *Neotyphodium* species with genus *Epichloë*. Mycologia *106*, 202-215.

Leyronas, C., and Raynal, G. (2008). Role of fungal ascospores in the infection of orchardgrass (*Dactylis glomerata*) by *Epichloë typhina* agent of choke disease. J Plant Pathol *90*, 15-21.

Liu, G., Casqueiro, J., Banuelos, O., Cardoza, R.E., Gutierrez, S., and Martin, J.F. (2001). Targeted inactivation of the *mecB* gene, encoding cystathionine-gamma-lyase, shows that the reverse transsulfuration pathway is required for high-level cephalosporin biosynthesis in *Acremonium chrysogenum* C10 but not for methionine induction of the cephalosporin genes. J Bacteriol *183*, 1765-1772.

Liu, X.Y., and Walsh, C.T. (2009). Cyclopiazonic acid biosynthesis in *Aspergillus* sp.: characterization of a reductase-like R* domain in cyclopiazonate synthetase that forms and releases cyclo-acetoacetyl-L-tryptophan. Biochemistry (Wash) *48*, 8746-8757.

Lohman, J.R., Ma, M., Cuff, M.E., Bigelow, L., Bearden, J., Babnigg, G., Joachimiak, A., Phillips, G.N., Jr., and Shen, B. (2014). The crystal structure of BlmI as a model for

nonribosomal peptide synthetase peptidyl carrier proteins. Proteins: Struct Funct Bioinform *82*, 1210-1218.

Louie, G.V., Baiga, T.J., Bowman, M.E., Koeduka, T., Taylor, J.H., Spassova, S.M., Pichersky, E., and Noel, J.P. (2007). Structure and reaction mechanism of basil eugenol synthase. PLoS One *2*, e993.

Lukito, Y., Chujo, T., and Scott, B. (2015). Molecular and cellular analysis of the pH response transcription factor PacC in the fungal symbiont *Epichloë festucae*. Fungal Genet Biol *85*, 25-37.

Marahiel, M.A., Stachelhaus, T., and Mootz, H.D. (1997). Modular peptide synthetases involved in nonribosomal peptide synthesis. Chemical Reviews *97*, 2651-2673.

May, J.J., Kessler, N., Marahiel, M.A., and Stubbs, M.T. (2002). Crystal structure of DhbE, an archetype for aryl acid activating domains of modular nonribosomal peptide synthetases. Proc Natl Acad Sci USA *99*, 12120-12125.

McMillan, L.K., Carr, R.L., Young, C.A., Astin, J.W., Lowe, R.G.T., Parker, E.J., Jameson, G.B., Finch, S.C., Miles, C.O., McManus, O.B.*, 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.

Miller, J.H. (1972). Experiments in molecular genetics (New York: Cold Spring Harbor Laboratory).

Min, T.P., Kasahara, H., Bedgar, D.L., Youn, B.Y., Lawrence, P.K., Gang, D.R., Halls, S.C., Park, H.J., Hilsenbeck, J.L., Davin, L.B.*, et al.* (2003). Crystal structures of pinoresinol-lariciresinol and phenylcoumaran benzylic ether reductases and their relationship to isoflavone reductases. J Biol Chem *278*, 50714-50723.

Minin, V.N., Dorman, K.S., Fang, F., and Suchard, M.A. (2005). Dual multiple change-point model leads to more accurate recombination detection. Bioinformatics *21*, 3034-3042.

Namiki, F., Matsunaga, M., Okuda, M., Inoue, I., Nishi, K., Fujita, Y., and 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.

Oide, S., Krasnoff, S.B., Gibson, D.M., and Turgeon, B.G. (2007). Intracellular siderophores are essential for ascomycete sexual development in heterothallic Cochliobolus heterostrophus and homothallic Gibberella zeae. Eukaryot Cell *6*, 1339-1353.

Oliver, R.P., Roberts, I.N., Harling, R., Kenyon, L., Punt, P.J., Dingemanse, M.A., and Vandenhondel, C. (1987). Transformation of *Fulvia fulva*, a fungal pathogen of tomato, to hygromycin-B resistance. Curr Genet *12*, 231-233.

Pan, J., Bhardwaj, M., Faulkner, J.R., Nagabhyru, P., Charlton, N.D., Higashi, R.M., Miller, A.-F., Young, C.A., Grossman, R.B., and Schardl, C.L. (2014a). Ether bridge formation in loline alkaloid biosynthesis. Phytochemistry *98*, 60-68.

Pan, J., Bhardwaj, M., Nagabhyru, P., Grossman, R.B., and Schardl, C.L. (2014b). Enzymes from fungal and plant origin required for chemical diversification of insecticidal loline alkaloids in grass-*Epichloë* symbiota. PloS one *9*, e115590.

Panaccione, D.G., Johnson, R.D., Wang, J.H., Young, C.A., Damrongkool, P., Scott, B., and Schardl, C.L. (2001). Elimination of ergovaline from a grass-*Neotyphodium* endophyte symbiosis by genetic modification of the endophyte. Proc Natl Acad Sci USA *98*, 12820-12825.

Pavela-Vrancic, M., Van Liempt, H., Pfeifer, E., Freist, W., and Von Dohren, H. (1994). Nucleotide binding by multienzyme peptide synthetases. Eur J Biochem *220*, 535-542.

Prieto, C., Garcia-Estrada, C., Lorenzana, D., and Francisco Martin, J. (2012). NRPSsp: non-ribosomal peptide synthase substrate predictor. Bioinformatics *28*, 426-427.

Rasmussen, S., Lane, G.A., Mace, W., Parsons, A.J., Fraser, K., and Xue, H. (2012). The use of genomics and metabolomics methods to quantify fungal endosymbionts and alkaloids in grasses. Methods Mol Biol *860*, 213-226.

Rausch, C., Weber, T., Kohlbacher, O., Wohlleben, W., and Huson, D.H. (2005). Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). Nucleic Acids Res *33*, 5799-5808.

Reuter, K., Mofid, M.R., Marahiel, M.A., and Ficner, R. (1999). Crystal structure of the surfactin synthetase-activating enzyme Sfp: a prototype of the 4'-phosphopantetheinyl transferase superfamily. EMBO J *18*, 6823-6831.

Roettig, M., Medema, M.H., Blin, K., Weber, T., Rausch, C., and Kohlbacher, O. (2011). NRPSpredictor2-a web server for predicting NRPS adenylation domain specificity. Nucleic Acids Res *39*, W362-W367.

Rowan, D.D., Dymock, J.J., and Brimble, M.A. (1990). Effect of fungal metabolite peramine and analogs on feeding and development of argentine stem weevil (*Listronotus bonariensis*). J Chem Ecol *16*, 1683-1695.

Saikia, S., Parker, E.J., Koulman, A., and Scott, B. (2006). Four gene products are required for the fungal synthesis of the indole-diterpene, paspaline. FEBS Lett *580*, 1625-1630.

Saikia, S., Takemoto, D., Tapper, B.A., Lane, G.A., Fraser, K., and Scott, B. (2012). Functional analysis of an indole-diterpene gene cluster for lolitrem B biosynthesis in the grass endosymbiont *Epichloë festucae*. FEBS Lett *586*, 2563-2569.

Schardl, C.L. (2010). The epichloae, symbionts of the grass subfamily *Pooideae*. Annals of the Missouri Botanical Garden *97*, 646-665.

Schardl, C.L., Grossman, R.B., Nagabhyru, P., Faulkner, J.R., and Mallik, U.P. (2007). Loline alkaloids: Currencies of mutualism. Phytochemistry *68*, 980-996.

Schardl, C.L., and Leuchtmann, A. (2005). The *Epichloë* endophytes of grasses and the symbiotic continuum. In The fungal community: its organization and role in the ecosystem, J. Dighton, J.F. White, and P. Oudemans, eds. (Boca Raton, Florida: CRC Press), pp. 475-503.

Schardl, C.L., Young, C.A., Faulkner, J.R., Florea, S., and Pan, J. (2012). Chemotypic diversity of epichloae, fungal symbionts of grasses. Fungal Ecol *5*, 331-344.

Schardl, C.L., Young, C.A., Hesse, U., Amyotte, S.G., Andreeva, K., Calie, P.J., Fleetwood, D.J., Haws, D.C., Moore, N., Oeser, B.*, et al.* (2013a). Plant-symbiotic fungi as chemical engineers: multi-genome analysis of the Clavicipitaceae reveals dynamics of alkaloid loci. PLoS Genet *9*.

Schardl, C.L., Young, C.A., Pan, J., Florea, S., Takach, J.E., Panaccione, D.G., Farman, M.L., Webb, J.S., Jaromczyk, J., Charlton, N.D.*, et al.* (2013b). Currencies of mutualisms: sources of alkaloid genes in vertically transmitted epichloae. Toxins *5*, 1064-1088.

Schardl, C.L., Young, C.A., Moore, N., Krom, N., Dupont, P.-Y., Pan, J., Florea, S., Webb, J.S., Jaromczyk, J., Jaromczyk, J.W., Cox, M.P., and Farman, M.L. (2014). Genomes of plant-associated Clavicipitaceae. Fungi *70*, 291-327.

Scharf, D.H., Heinekamp, T., and Brakhage, A.A. (2014). Human and plant fungal pathogens: the role of secondary metabolites. PLoS Path *10*.

Schneider, K., Hovel, K., Witzel, K., Hamberger, B., Schomburg, D., Kombrink, E., and Stuible, H.P. (2003). The substrate specificity-determining amino acid code of 4-coumarate : CoA ligase. Proc Natl Acad Sci USA *100*, 8601-8606.

Schneider, P., Bouhired, S., and Hoffmeister, D. (2008). Characterization of the atromentin biosynthesis genes and enzymes in the homobasidiomycete Tapinella panuoides. Fungal Genet Biol *45*, 1487-1496.

Schwarzer, D., Mootz, H.D., Linne, U., and Marahiel, M.A. (2002). Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases. Proc Natl Acad Sci USA *99*, 14083-14088.

Scott, B., Takemoto, D., Tanaka, A., Young, C., Bryant, M., and May, K. (2007). Functional analysis of the *Epichloë festucae*-perennial ryegrass symbiosis. Paper presented at: 6th International symposium on fungal endophytes of grasses New Zealand Grassland Association (Christchurch, New Zealand).

Scott, B., Young, C.A., Saikia, S., McMillan, L.K., Monahan, B.J., Koulman, A., Astin, J., Eaton, C.J., Bryant, A., Wrenn, R.E.*, et al.* (2013). Deletion and gene expression analyses define the paxilline biosynthetic gene cluster in *Penicillium paxilli*. Toxins *5*, 1422-1446.

Selosse, M.A., and Schardl, C.L. (2007). Fungal endophytes of grasses: hybrids rescued by vertical transmission? An evolutionary perspective. New Phytol *173*, 452-458.

Siegel, M.R., Latch, G.C.M., Bush, L.P., Fannin, F.F., Rowan, D.D., Tapper, B.A., Bacon, C.W., and Johnson, M.C. (1990). Fungal endophyte-infected grasses - alkaloid accumulation and aphid response. J Chem Ecol *16*, 3301-3315.

Sommer, S. (2005). The importance of immune gene variability (MHC) in evolutionary ecology and conservation. Frontiers in Zoology *2*, 16-16.

Spiering, M.J., Faulkner, J.R., Zhang, D.-X., Machado, C., Grossman, R.B., and Schardl, C.L. (2008). Role of the LolP cytochrome P450 monooxygenase in loline alkaloid biosynthesis. Fungal Genet Biol *45*, 1307-1314.

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

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

Stachelhaus, T., and Marahiel, M.A. (1995). Modular structure of genes encoding multifunctional peptide synthetases required for non-ribosomal peptide synthesis. FEMS Microbiol Lett *125*, 3-14.

Stachelhaus, T., Mootz, H.D., and Marahiel, M.A. (1999). The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. Chem Biol *6*, 493-505.

Stammers, D.K., Ren, J., Leslie, K., Nichols, C.E., Lamb, H.K., Cocklin, S., Dodds, A., and Hawkins, A.R. (2001). The structure of the negative transcriptional regulator NmrA reveals a structural superfamily which includes the short-chain dehydrogenase/reductases. EMBO J *20*, 6619-6626.

Tadych, M., Ambrose, K.V., Bergen, M.S., Belanger, F.C., and White, J.F. (2012). Taxonomic placement of *Epichloë poae* sp. *nov*. and horizontal dissemination to seedlings via conidia. Fungal Diversity *54*, 117-131.

Takemoto, D., Kamakura, S., Saikia, S., Becker, Y., Wrenn, R., Tanaka, A., Sumimoto, H., and Scott, B. (2011). Polarity proteins Bem1 and Cdc24 are components of the filamentous fungal NADPH oxidase complex. Proc Natl Acad Sci U S A *108*, 2861-2866.

Takemoto, D., Tanaka, A., and Scott, B. (2006). A p67(Phox)-like regulator is recruited to control hyphal branching in a fungal-grass mutualistic symbiosis. Plant Cell *18*, 2807-2821.

Tamura, K. (1992). Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C-content biases. Mol Biol Evol *9*, 678-687.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol *28*, 2731-2739.

Tan, Y.Y., Spiering, M.J., Scott, V., Lane, G.A., Christensen, M.J., and Schmid, J. (2001). In planta regulation of extension of an endophytic fungus and maintenance of high metabolic rates in its mycelium in the absence of apical extension. Appl Environ Microbiol *67*, 5377-5383.

Tanaka, A., Cartwright, G.M., Saikia, S., Kayano, Y., Takemoto, D., Kato, M., Tsuge, T., and Scott, B. (2013). ProA, a transcriptional regulator of fungal fruiting body development, regulates leaf hyphal network development in the *Epichloë festucae-Lolium perenne* symbiosis. Mol Microbiol *90*, 551-568.

Tanaka, A., Christensen, M.J., Takemoto, D., Park, P., and Scott, B. (2006). Reactive oxygen species play a role in regulating a fungus-perennial ryegrass mutualistic interaction. Plant Cell *18*, 1052-1066.

Tanaka, A., Takemoto, D., Chujo, T., and Scott, B. (2012). Fungal endophytes of grasses. Curr Opin Plant Biol *15*, 462-468.

Tanaka, A., Takemoto, D., Hyon, G.S., Park, P., and Scott, B. (2008). NoxA activation by the small GTPase RacA is required to maintain a mutualistic symbiotic association between *Epichloë festucae* and perennial ryegrass. Mol Microbiol *68*, 1165-1178.

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

Telfer, E. (2000). Regulation of paxilline biosynthesis in *Penicillium paxilli*. In Institute of Molecular Biosciences (Palmerston North, New Zealand: Massey University).

Tesicky, M., and Vinkler, M. (2015). Trans-species polymorphism in immune genes: general pattern or MHC-restricted phenomenon? J Immunol Res.

Tofern, B., Kaloga, M., Witte, L., Hartmann, T., and Eich, E. (1999). Occurrence of loline alkaloids in *Argyreia mollis* (Convolvulaceae). Phytochemistry *51*, 1177-1180.

Tsai, H.-F., Wang, H., Gebler, J.C., Poulter, C.D., and Schardl, C.L. (1995). The *Claviceps purpurea* gene encoding dimethylallyltryptophan synthase, the committed step for ergot alkaloid biosynthesis. Biochem Biophys Res Commun *216*, 119-125.

Tufar, P., Rahighi, S., Kraas, F.I., Kirchner, D.K., Loehr, F., Henrich, E., Koepke, J., Dikic, I., Guentert, P., Marahiel, M.A.*, et al.* (2014). Crystal structure of a PCP/Sfp complex reveals the structural basis for carrier protein posttranslational modification. Chem Biol *21*, 552-562.

Vanden Wymelenberg, A.J., Cullen, D., Spear, R.N., Schoenike, B., and Andrews, J.H. (1997). Expression of green fluorescent protein in *Aureobasidium pullulans* and quantification of the fungus on leaf surfaces. BioTechniques *23*, 686-690.

Velkov, T., Horne, J., Scanlon, Martin J., Capuano, B., Yuriev, E., and Lawen, A. (2011). Characterization of the N-methyltransferase activities of the multifunctional polypeptide cyclosporin synthetase. Chem Biol *18*, 464-475.

von Döhren, H. (2004). Biochemistry and general genetics of nonribosomal peptide synthetases in fungi. In Molecular Biotechnolgy of Fungal beta-Lactam Antibiotics

and Related Peptide Synthetases, A.A. Brakhage, ed. (Springer Berlin Heidelberg), pp. 217-264.

Walsh, C.T., Gehring, A.M., Weinreb, P.H., Quadri, L.E.N., and Flugel, R.S. (1997). Post-translational modification of polyketide and nonribosomal peptide synthases. Curr Opin Chem Biol *1*, 309-315.

Wang, H., Fewer, D.P., Holm, L., Rouhiainen, L., and Sivonen, K. (2014). Atlas of nonribosomal peptide and polyketide biosynthetic pathways reveals common occurrence of nonmodular enzymes. Proc Natl Acad Sci USA *111*, 9259-9264.

Wang, J., Machado, C., Panaccione, D.G., Tsai, H.-F., and Schardl, C.L. (2004). The determinant step in ergot alkaloid biosynthesis by an endophyte of perennial ryegrass. Fungal Genet Biol *41*, 189-198.

Wang, X.Q., He, X.Z., Lin, J.Q., Shao, H., Chang, Z.Z., and Dixon, R.A. (2006). Crystal structure of isoflavone reductase from alfalfa (*Medicago sativa* L.). J Mol Biol *358*, 1341-1352.

Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H.U., Bruccoleri, R., Lee, S.Y., Fischbach, M.A., Mueller, R., Wohlleben, W*., et al.* (2015). antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res *43*, W237-W243.

Wilson, R.C., and Doudna, J.A. (2013). Molecular Mechanisms of RNA Interference. Ann Rev Biophys *42*, 217-239.

Winston, F., Dollard, C., and Ricupero-Hovasse, S.L. (1995). Construction of a set of convenient *Saccharomyces cerevisiae* strains that are isogenic to S288C. Yeast *11*, 53-55.

Wisecaver, J.H., and Rokas, A. (2015). Fungal metabolic gene clusters-caravans traveling across genomes and environments. Frontiers in Microbiology *6*.

Wisecaver, J.H., Slot, J.C., and Rokas, A. (2014). The Evolution of Fungal Metabolic Pathways. PLoS Genet *10*.

Yeh, E., Kohli, R.M., Bruner, S.D., and Walsh, C.T. (2004). Type II thioesterase restores activity of a NRPS module stalled with an aminoacyl-S-enzyme that cannot be elongated. ChemBioChem *5*, 1290-1293.

Yelton, M.M., Timberlake, W.E., and Hondel, C.A. (1985). A cosmid for selecting genes by complementation in *Aspergillus nidulans*: selection of the developmentally regulated yA locus. Proc Natl Acad Sci U S A *82*, 834-838.

Young, C., Itoh, Y., Johnson, R., Garthwaite, I., Miles, C.O., Munday-Finch, S.C., and Scott, B. (1998). Paxilline-negative mutants of *Penicillium paxilli* generated by heterologous and homologous plasmid integration. Curr Genet *33*, 368-377.

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

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

Young, C.A., Felitti, S., Shields, K., Spangenberg, G., Johnson, R.D., Bryan, G.T., Saikia, S., and Scott, B. (2006). A complex gene cluster for indole-diterpene biosynthesis in the grass endophyte *Neotyphodium lolii*. Fungal Genet Biol *43*, 679-693.

Young, C.A., Hume, D.E., and McCulley, R.L. (2013). Fungal endophytes of tall fescue and perennial ryegrass: pasture friend or foe? J Anim Sci *91*, 2379-2394.

Young, C.A., Schardl, C.L., Panaccione, D.G., Florea, S., Takach, J.E., Charlton, N.D., Moore, N., Webb, J.S., and Jaromczyk, J. (2015). Genetics, genomics and evolution of ergot alkaloid diversity. Toxins *7*, 1273-1302.

Young, C.A., Tapper, B.A., May, K., Moon, C.D., Schardl, C.L., and Scott, B. (2009). Indole-diterpene biosynthetic capability of *Epichloë* endophytes as predicted by *LTM* gene analysis. Appl Environ Microbiol *75*, 2200-2211.

Youssar, L., Gruning, B.A., Erxleben, A., Gunther, S., and Huttel, W. (2012). Genome sequence of the fungus *Glarea lozoyensis*: the first genome sequence of a species from the Helotiaceae family. Eukaryot Cell *11*, 250.

Zabala, A.O., Chooi, Y.-H., Choi, M.S., Lin, H.-C., and Tang, Y. (2014). Fungal polyketide synthase product chain-length control by partnering thiohydrolase. ACS Chemical Biology *9*, 1576-1586.

Zhou, Z., Lai, J.R., and Walsh, C.T. (2006). Interdomain communication between the thiolation and thioesterase domains of EntF explored by combinatorial mutagenesis and selection. Chem Biol *13*, 869-879.