IDENTIFICATION OF PUTATIVE DOTHISTROMIN BIOSYNTHETIC GENES.

A thesis presented in partial fulfilment of the requirements
for the degree of Master of Science in Molecular Biology
at Massey University, Palmerston North, New Zealand.

Brendon Joseph Monahan

1998

ABSTRACT.

Dothistromin is a polyketide-derived mycotoxin produced by the *Pinus* pathogen *Dothistroma pini*, and is thought to be important in the development of the necrotic disease *Dothistroma* needle blight. Targeted disruption of dothistromin biosynthetic genes will allow the direct assessment of the role of the toxin in *D. pini* pathogenicity. Dothistromin displays structural and biochemical similarities to the aflatoxins (AF) and sterigmatocystin (ST) which are produced by various *Aspergillus* species. In our laboratory, knowledge from the well characterised ST/AF pathway is being used to isolate and characterise genes likely to be involved in dothistromin production.

The *D. pini* lambda clone, λCGV1, was isolated from a *D. pini* genomic library by heterologous hybridisation with a fragment of the *Aspergillus parasiticus ver1* gene (Gillman, 1996). In this study, the complete nucleotide sequence of λCGV1 was determined. Analysis revealed that five genes are located within the 13.3 kb genomic region sequenced. Three of these genes (*dkr1, dox1* and *dte1*) display strong similarities to genes contained within the ST/AF biosynthetic gene clusters. The *dtp1* gene, located between *dox1* and *dte1*, shows similarities to transmembrane efflux pumps and is proposed to be a dothistromin toxin pump. The *ddh1* gene, located upstream of *dkr1*, shows similarities to bacterial dehydrogenases. However, the *ddh1* coding sequence contains a premature stop codon (encoding a product of 63 amino acids), indicating that the product may be non-functional.

Expression analysis of each gene identified in this study confirmed that *dkr1*, *dox1*, *dte1* and *dtp1* are expressed. However, no obvious expression was detected for the *ddh1* gene. Southern blot analysis confirmed the genomic clustering of the genes and indicated that a single copy of each gene was present in the *D. pini* genome.

Due to the biogenetic relationship between dothistromin and ST/AF biosynthesis, and because genes identified in this study show similarities to genes involved in ST/AF production, it is thought that these genes are likely to be involved in dothistromin biosynthesis and constitute part of a dothistromin biosynthetic gene cluster.

ACKNOWLEDGEMENTS.

My sincerest thanks goes to my supervisor Rosie Bradshaw for her support and encouragement throughout my project. A big thanks also to Peter Lockhart for his expert help with the phylogenetic analysis and to Bill Jones (and lab) for much needed guidance with the ELISA work, and also for some very informative chats. I would also like to thank the Holt Forest Trust for financial assistance.

I am very appreciative for the support I have received in and out of the lab. A big thanks to Paul, Linda, Dianne, Anita, Carmel, Bran, Tania, Lisa, Mike and Tash for answering all of my many questions when I started (and beyond). Answering many more questions has been Carolyn, and I thank you for all your effort and time. And to the present lab members (Bek, Beccy and Seth) thanks for your support and best of luck with your research.

Special thanks to all my flatmates, especially Pig and Dave, its been a good five and a bit years and cheers for the many many laughs. Thanks also to anyone that has had a beer(s) and a chat with me over at Wharerata (or anywhere else for that matter) over the last few years, especially Austen, David and a couple of Irish people.

Many thanks and best wishes to Helen, Kirsty and Amanda for all the support and company at lunch. A big thanks must go to Austen for everything, keep up the great work. Also a big thank-you to Bek, for all your help and encouragement, especially over the last few months.

Finally, I have my family to thank for many many things. Thanks to all, especially mum, for your love, encouragement and support. Your effort is a huge factor in the completion of this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES.

1. INTRODUCTION.

1.1 *DOTHISTROMA* NEEDLE BLIGHT.

The filamentous fungus *Dothistroma pini* Hulbary is a needle pathogen of *Pinus radiata* and other *Pinus* species, producing a necrotic disease known as *Dothistroma* needle blight. This disease is characterised by red bands through necrotic lesions on the needle and can cause premature needle cast (initially at the base of the crown) and consequent reduction of photosynthesis and wood yield. In extreme cases, *D. pini* infection can cause death of the infected tree (Gadgil, 1984).

Dothistroma needle blight has been recorded in Europe, South and East Africa, North and South America, and Australia and was first identified in New Zealand (Kinleith Forest) in 1962 (Dick, 1989). The disease is distributed throughout New Zealand except in the northern extremity of the North Island (Gadgil, 1984). From a forestry and economic viewpoint, *P. radiata* is the most important susceptible species and in several parts of the world *D. pini* infection has caused serious defoliation of *P. radiata* plantations.

The forest industry in New Zealand is an expanding and important resource industry contributing 6.9% to New Zealand's gross domestic product in 1996. Overseas earnings from forestry imports are increasing and projected to do so for the next twenty years. New Zealand has 1.5 million hectares of planted production forest of which 90.5% (1.3 million hectares) is *P. radiata*, contributing 34.9% of the world's radiata forest estate (New Zealand Forestry Facts and Figures 1996). *D. pini* is the major pathogen of *P. radiata* and infects predominantly younger trees (less than twenty years). With the majority of *P. radiata* trees in the New Zealand forests being less than twenty years of age, *D. pini* poses a significant threat to the forest industry.

1.2 *D. PINI* INFECTION OF *P. RADIATA*.

Dothistroma pini (also known as *Dothistroma septospora*) is of the order Dothideales in the class Ascomycotina, and is the anamorphic (asexual) form of *Mycosphaerella pini*

(also known as *Scirrhia pini*). The sexual form has not been observed in New Zealand (Gadgil 1967, 1984; Evans, 1984).

A film of water on an infected needle is the source of rain splash dispersal of *D. pini* conidia (asexual spores). Conidia-containing droplets drip off the infected needle and on contact with branches and/or the ground spores are released into the air and subsequent spread of infection can occur. However, spores released in this manner do not usually travel far and infection is generally from branch to branch, neighbour to neighbour. Widespread dispersal is usually the result of conidia being caught up in mist or cloud (Gadgil, 1984).

Dispersed conidia attach to the surface of the pine needle and most germinate within three days, with each conidium producing one to three germ tubes. After seven to ten days an abundant growth of mycelium over the surface of the needle is observed, as is production of secondary conidia (Gadgil, 1967). Appressorium-like structures form, just covering the stomatal openings, and penetration of hyphae into the needle occurs through the stomata. Once inside the needle, inter- and intra-cellular hyphal growth proceeds (Gadgil, 1967, 1984). The physical appearance of a *D. pini* infected needle is a yellow area (due to chlorosis) which develops into (and occasionally surrounds) the characteristic brick red band. In these red bands small black fruiting bodies (stromata) develop which produce conidia and the infection cycle, depending on the correct conditions, continues (Gadgil, 1984).

A number of factors influence the establishment and severity of infection; the duration of the needle wetness period, the number of conidia landing on the needle surface, and the temperature. Severe infection occurs when the needle surface remains wet for at least ten hours, with several thousand conidia present and a mean daily temperature of 16-18°C (Gadgil, 1984).

Mature *P. radiata* trees (greater than 15-20 years) generally exhibit a significant level of resistance to *D. pini* infection, compared to the susceptibility of younger trees. Mature tree resistance is hereditary and is transferred by grafting. Differences in the size and state of stomata and chemical differences in the composition of wax around the stomata

between young and mature trees are considered important to the mode of resistance, but overall the exact nature of mature tree resistance is not known (Franich and Wells, 1977; Franich *et al.*, 1977; Franich *et al.*, 1978).

1.3 CURRENT INFECTION CONTROL METHODS: WHY THE NEED FOR OTHERS.

The major strategy for control of *D. pini* infection is the aerial application of copper fungicides. Copper fungicide inhibits germination of *D. pini* conidia and reduces germ tube growth, thus preventing severe infection and spread (Franich, 1988). Aerial spraying was initiated in the Kinleith forest in 1966. Dick (1989) reported that the total cost of this programme in New Zealand up to the end of the 1988-89 summer was $18.4 million (1988 dollars).

The aerial spraying programme involves aerial surveillance in mid-winter from which the average percentage crown infection of stands is assessed. Infection scores less than 15% require no copper fungicide spray treatment, 15-30% require one spray application, and average crown infection levels greater than 30% require two spray applications, one administered in early summer and the other in late summer. Pruning of lower branches to remove the microclimate necessary for infection is often carried out in conjunction with spraying to control disease spread (Gadgil, 1984, Dick, 1989).

Another infection control strategy is the *D. pini* resistant *P. radiata* breeding programme which was established by the New Zealand Forest Research Institute (NZFRI). This programme has produced trees with increased resistance to *D. pini* infection, however, a mean infection reduction of only 15% is observed. As a consequence of vigorous selection for *D. pini* resistance, a reduction in growth rate of resistance breed trees is also observed, so different breeds are used in areas without a significant *D. pini* problem (Carson and Carson, 1989, 1991).

A detailed study investigating the genetic diversity of *D. pini* in New Zealand was undertaken by Hirst (1997). The main technique used in this investigation was random amplified polymorphic DNA (RAPD) analysis. The study concluded that the genetic diversity of *D. pini* in New Zealand was very low and that all New Zealand isolates

examined appeared to be of a single strain. These results imply that the resistance breeding program has generated resistance to the one *D. pini* strain. Thus different *D. pini* strains could be more virulent on trees which currently show some degree of *D. pini* resistance, and introduction of these strain(s) into New Zealand could be quite destructive to the forestry industry.

Research is currently being performed in our laboratory to determine whether the current *D. pini* strain in New Zealand is asexual, or is sexual but only of a single mating type. Introduction into New Zealand of *M. pini* or a strain of the opposite mating type would allow sexual recombination to occur which would cause greater genetic variation of the fungus and could lead to increased virulence (Wang, 1997).

The above risks, coupled with current domestic and global infection levels and associated costs, highlights the need for comprehensive protection of commercial radiata pine from *Dothistroma* needle blight. This requires detailed research into the pathogenesis and biology of *D. pini*. A molecular based investigation into the biosynthesis of the toxin dothistromin, which *D. pini* produces, is the focus of this study.

1.4 DOTHISTROMIN.

Gadgil (1967) first noted the possibility of toxin action in his characterisation of *D. pini* infection of *P. radiata*. This resulted from observations of mesophyll disruption well in advance of fungal hyphae and of dead cells adjacent to areas colonised by *D. pini* hyphae. Further studies (Bassett and Buchanan, 1970; Gallagher and Hodges, 1972; Danks and Hodges, 1974; Shaw *et al.*, 1978; Franich, 1981; Stoessl *et al.*, 1990) reported the isolation and characterisation of the structural, toxic and biochemical properties of a difuroanthraquinone mycotoxin named dothistromin.

Dothistromin is the major metabolic by-product of *D. pini* and was shown to play a key role in pathogenesis by experiments in which *P. radiata* needles injected with dothistromin developed needle blight symptoms. Control injections of acetone alone failed to produce such symptoms (Shain and Franich, 1981). Symptoms were induced with dothistromin concentrations (10-100ng) substantially less than concentrations detected in naturally occurring lesions (1-10mg) (Shain and Franich, 1981). However, the

exact role of dothistromin in pathogenesis has not been determined. This requires the generation and analysis of *D. pini* mutants defective in dothistromin production. It is therefore important to characterise dothistromin biosynthesis at the molecular level. This would allow targeted disruption of appropriate genes to be undertaken and possibly the generation of dothistromin-minus mutants.

1.4.1 Disruption of Toxin Biosynthesis.

Many studies have sought to disrupt the synthesis of a toxin implicated in pathogenesis. Cerato-ulmin (CU) is a low molecular weight toxin produced by the Dutch elm disease (DED) causing pathogen *Ophiostoma ulmi*. Like dothistromin, *in vitro* tests using purified CU induced DED like symptoms when administered to elm saplings. To directly test the role of CU in pathogenesis, the *O. ulmi* CU gene was cloned (Bowden *et al.*, 1994) and a CU-minus mutant created by transformation-mediated gene disruption. However, this mutant retained full pathogenicity (Bowden *et al.*, 1996). The fungus *Cercospora kikuchii* is a soybean pathogen and produces a light sensitive polyketide toxin, cercosporin. Once again *in vitro* tests implicated the toxin in having a direct role in pathogenesis. In this case, however, cercosporin minus mutants were shown to be non-pathogenic to soybean plants (Upchurch *et al.*, 1991), confirming that cercosporin is essential for pathogenicity.

To directly test the role of dothistromin in needle blight disease it is necessary to produce dothistromin-minus mutants. Transformation mediated gene disruption is the most suitable technique for targeted disruption of key biosynthetic genes because the multinucleate nature of *D. pini* conidia makes conventional *uv*-mutagenesis techniques difficult to employ.

Elucidation of the role of dothistromin in pathogenesis is vital in the overall aim to prevent *Dothistroma* needle blight. If dothistromin-minus mutants are non-pathogenic a number of possible options arise to protect trees from infection. Stable dothistromin-minus mutants could be released into the wild with the aim of the non-pathogenic strain out-competing the wild type strain. Another option is to produce transgenic trees which express antibodies to dothistromin, or other molecules which neutralises toxin action. This latter approach alongside investigation into the role of dothistromin in pathogenesis

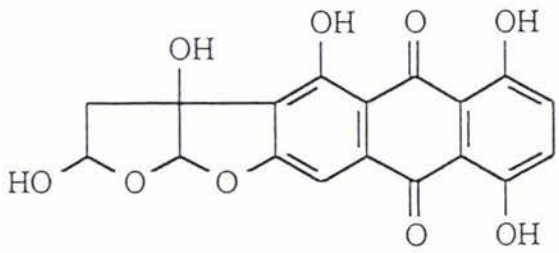
by molecular immunology is the focus of a Hort Research group in Palmerston North, lead by Bill Jones (formally lead by Paul Reynolds). This group has established an enzyme-linked immunosorbent assay (ELISA) for dothistromin detection (Jones *et al.*, 1993) and recently the production of anti-idiotypic monoclonal antibodies that mimic dothistromin (Jones *et al.*, 1998) which are being used to detect dothistromin binding receptors.

Dothistromin is a polyketide-derived phytotoxin and contains a furobenzofuran moiety. This is a common structural feature of the aflatoxins, sterigmatocystins and the versicolorins (Figure 1.1) which are produced by some *Aspergillus* species (Gallagher and Hodges, 1972). The furan ring structure is associated with the toxicity of these compounds. Dothistromin is also produced by some *Cercospora* species and biochemical studies with both the peanut pathogen *C. arachidicola* and *D. pini* identified compounds which are known intermediates in the sterigmatocystin and aflatoxin biosynthetic pathway (Figure 1.2) (Danks and Hodges, 1974; Assante *et al.*, 1977; Stoessl, 1984). These similarities indicate a biogenetic relationship between dothistromin and aflatoxin biosynthesis (Steyn, 1980). If sufficient similarity exists, dothistromin biosynthetic genes can be targeted and isolated using knowledge from the aflatoxin biosynthetic pathway (for a review of methods for cloning heterologous genes see Agnan *et al.*, 1997). It is therefore important to understand alfatoxin biosynthesis and regulation.

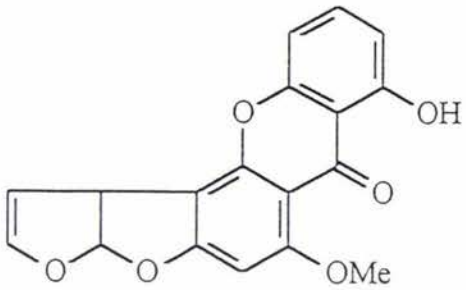
1.5 AFLATOXINS.

Aflatoxins (AF) are a group of biologically active secondary metabolites with similar chemical structures, which are predominantly produced by certain strains of *Aspergillus flavus* and *A. parasiticus*. Many other *Aspergillus* species, most notably *A. nidulans*, produce AF intermediates (up to the highly toxic sterigmatocystin [ST]) but do not produce AF, thus indicating a metabolic block in AF biosynthesis.

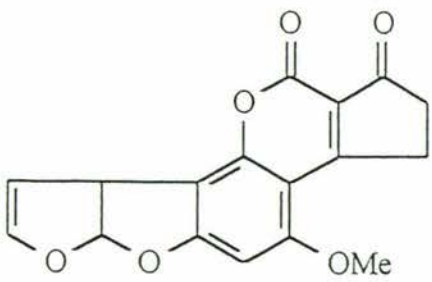
Of the aflatoxins, aflatoxin B_1 (AFB_1) is the most abundant and toxic. AFB_1 is the most potent hepatocarcinogen known for the rat and rainbow trout (Steyn, 1980). Worldwide, AFs are considered a major public health and economic problem as the ubiquitous fungi which produce them are known contaminants of important agricultural commodities such as corn, peanuts and cottonseed, but also of meat, eggs and milk from animals that have consumed contaminated feed. Ingestion of food or feed highly contaminated with AF can



DOTHISTROMIN.



STERIGMATOCYSTIN.



AFLATOXIN B_1.

Figure 1.1 Structures of Dothistromin, Sterigmatocystin, and Aflatoxin B1.

lead to acute toxicity including hepatotoxicity, teratogenicity, immunotoxicity and even death (Trail *et al.*, 1995). Studies investigating the toxicity of dothistromin to animals concluded that dothistromin is a mutagen and a potential carcinogen, but not as potent as AFB_1 and does not pose a significant risk to forestry workers (Elliott *et al.*, 1989).

Since 1960 when AFB_1 was identified as the cause of Turkey-X disease, study into numerous aspects of AFs has been immense. The goal of AF research is to unravel the circumstances of the ST/AFB_1 biosynthetic pathway and its regulation to allow the control or eradication of AFB_1 production. This has led to detailed biochemical analysis of AF production, supplemented by ever-increasing detailed molecular data (Woloshuk and Prieto, 1998; Minto and Townsend, 1997; Trail *et al.*, 1995 for reviews; Figure 1.2 for proposed pathway).

1.6 MOLECULAR BIOLOGY OF AFLATOXIN BIOSYNTHESIS.

Aflatoxin biosynthesis is initiated by the production of a six-carbon fatty acid, hexanoate, catalysed by a fatty acid synthetase (FAS). Hexanoate serves as a primer for extension by a polyketide synthase (PKS) to generate the first stable intermediate, norsolorinic acid. This is in contrast to most polyketide synthases which utilise acetate as the starter unit, and was a controversial theory until both *PKS* and *FAS* genes were isolated from *A. parasiticus* (Chang *et al.*, 1995) and *A. nidulans* (Yu and Leonard, 1995; Brown *et al.*, 1996). This serves to illustrate the power of molecular genetics in the characterisation and elucidation of biosynthetic pathways. A summary of the proposed ST/AFB_1 biosynthetic pathway is shown in Figure 1.2.

1.6.1 Isolation of Aflatoxin/Sterigmatocystin Biosynthetic Genes.

Strategies used for isolation of ST/AFB_1 biosynthetic genes include genetic complementation, reverse genetics, and polymerase chain reaction (PCR)-based methods to isolate homologs in other systems.

Genetic complementation isolates a gene by its ability to restore wild-type activity to a characterised defective mutant. For example, the *ver1* gene of *A. parasiticus* was isolated when a cosmid containing the gene (as part of a cosmid library constructed from wild-type *A. parasiticus* genomic DNA) was transformed into, and restored wild-type activity

to, the *A. parasiticus* mutant CS10 (Skory *et al.*, 1992). This mutant was defective in the conversion of versicolorin A (VA) to ST, thus accumulated the yellow intermediate VA (Lee *et al.*, 1975). This method has been used to identify other genes involved in the pathway such as *nor-1* (Chang *et al.*, 1992), *ord1* (Prieto *et al.*, 1996; Prieto and Woloshuk, 1997), the *avf1* locus (Prieto *et al.*, 1996), and the regulatory gene *aflR* (Payne *et al.*, 1993).

Reverse genetics is a process where the selected protein is purified and used to provide information that leads back to the isolation of its cognate gene. This is achieved by two main methods. The amino acid sequence of the protein is determined and used in the design of oligonucleotide primers for use in a PCR reaction, with the resulting product used as a probe to isolate the gene by screening a genomic or cDNA library. Alternatively, antibodies to the purified protein are produced and used to isolate the gene from a cDNA expression library in *E. coli*. Reverse genetics methods have led to the isolation of pathway genes such as *omt-A* (Yu *et al.*, 1993), *vbs* (Silva *et al.*, 1996) and *norA* (Cary *et al.*, 1996).

Isolation of the *aflR* gene in *A. nidulans* (Yu *et al.*, 1996) is an example of using degenerate primers based on deduced amino acid sequences of previously described genes in PCR reactions (in a similar manner to above). Putative *ver*-homologs have also been isolated in this manner from other *Aspergillus* species (Kusumoto *et al.*, 1996).

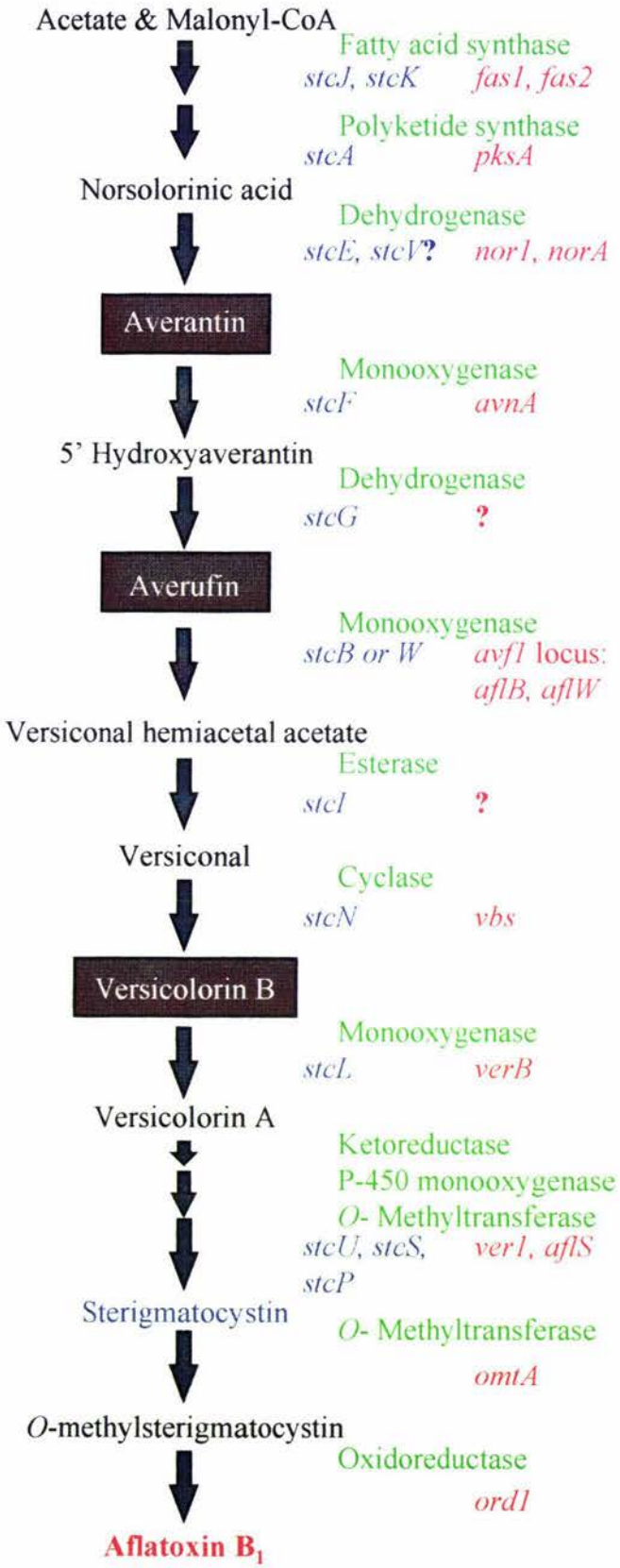
In all cases, gene function was confirmed by nucleotide sequencing and gene disruption (recombinational inactivation).

The factor which has had the greatest influence in the isolation of biosynthetic genes is that genes required for ST/AFB_1 biosynthesis are clustered. This has enabled the rapid identification of genes by targeting regions neighbouring previously described genes. The present status with regards to ST/AFB_1 genes and their proposed function is illustrated in Figure 1.2.

Figure 1.2 The Proposed Sterigmatocystin/Aflatoxin B1 Biosynthetic Pathway and Common Intermediates Found in Dothistromin-Producing Cultures.

The ST/AFB_1 biosynthetic pathway with the major intermediates is shown. Compounds which have also been identified in dothistromin-producing *D. pini* and *Cercospora smilacis* cultures are boxed (Danks and Hodges, 1974; Assante *et al.*, 1977). The predicted enzyme activity required for each step is indicated, as are the names of genes shown (or believed to) encode the respective product. Sterigmatocystin and genes involved in ST biosynthesis in *A. nidulans* are in blue; Aflatoxin B_1 and genes involved in AFB_1 biosynthesis in *A. parasiticus* and *A. flavus* are in red. Question marks (?) indicate steps that no genes have been assigned to (or designation is uncertain). Figure adapted from Yu *et al.* (1995), Trail *et al.* (1995), Brown *et al.* (1996), Prieto *et al.* (1996), Minto and Townsend (1997) and Woloshuk and Prieto (1998).

Proposed Sterigmatocystin / Aflatoxin B_1 Biosynthetic Pathway.



1.6.2 The ST/AFB$_1$ Biosynthetic Gene Cluster.

It is estimated that 16-17 enzyme-catalysed steps are required to complete AF biosynthesis and parasexual genetic studies established the likely clustering of at least some of the genes involved (Bradshaw *et. al.*, 1983). This has been confirmed through recent molecular studies in *A. nidulans*, *A. parasiticus* and *A. flavus* (Brown *et al.*, 1996; Yu *et al.*, 1995; Prieto *et al.*, 1996).

Brown *et al.* (1996) reported the characterisation of a 60 kb region in the *A. nidulans* genome that contains 25 coregulated genes which define most, if not all, of the enzymatic activities required for ST biosynthesis. Expression analysis established that all 25 transcripts were present in cultures under ST-inducing conditions but absent in cultures under non-ST inducing conditions. The two borders of the cluster were defined by transcripts located 5` (~6.5 kb from *stcA*) and 3` (~12 kb from *stcX*) to the cluster that, unlike the coregulated *stc* transcripts, were present in cultures at all time points. Of the 25 open reading frames (ORFs), 20 contain significant identity to previously described sequences, of which six; *stcA* (PKS, Yu and Leonard, 1995), *stcS* (P-450 monooxygenase, Keller *et al.*, 1995), *stcU* (ketoreductase, Keller *et al.*, 1994), *aflR* (transcription factor, Yu *et al.*, 1996), *stcP* (*O*-methyltransferase, Kelkar *et al.*, 1996) and *stcL* (P-450 monooxygenase, Kelkar *et al.*, 1997) have been shown to be required for ST biosynthesis in *A. nidulans* and a further three (*stcJ*, *stcK*, and *stcN*) have been implicated (Brown *et. al.*, 1996; Figure 1.2 and Figure 1.3).

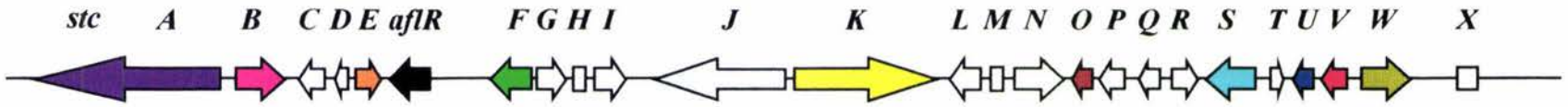
Arrangement of *A. parasiticus* and *A. flavus* cosmid and lambda clones containing previously described AFB$_1$ biosynthetic genes confirmed that genes in each system were also clustered within a 60 kb region (Yu *et al.*, 1995). Comparison of the *A. parasiticus* and *A. flavus* clusters showed that the gene order and distances between specific genes were similar. The order of the genes also corresponded with the order of respective pathway enzyme activities. However, more recentdiscovery of additional genes in gaps has disrupted this order somewhat. A study in *A. flavus* showed that all the genes required for AFB$_1$ biosynthesis are located within a 90 kb genomic DNA region (Prieto *et al.*, 1996). Presently 18 genes (over 75 kb) have been identified as part of the AFB$_1$

Figure 1.3 Sterigmatocystin and Aflatoxin B1 Biosynthetic Gene Clusters.

(a) The *A. nidulans* ST cluster (stc). **(b)** The *A. flavus*/*A. parasiticus* AFB_1 cluster. Arrows indicate genes and direction of transcription (*stcH*, *stcM* and *stcX* direction is not known). Common gene colour between clusters indicates significant sequence identity between genes, or gene products thought to be involved at the same step. The ST cluster is shown to scale whereas the AFB_1 cluster is only approximate. Figure adapted from Yu *et al.* (1995), Brown *et al.* (1996), Minto and Townsend (1997) and Woloshuk and Prieto (1998).

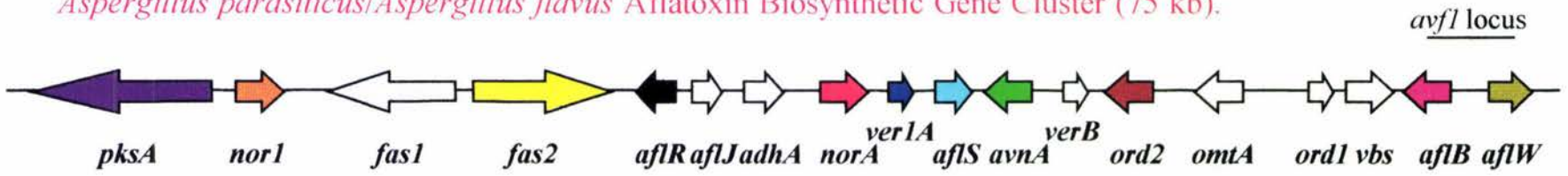
a)

Aspergillus nidulans Sterigmatocystin Biosynthetic Gene Cluster (60 kb).



b)

Aspergillus parasiticus/*Aspergillus flavus* Aflatoxin Biosynthetic Gene Cluster (75 kb).



biosynthetic cluster (Figure 1.3; reviewed in Minto and Townsend, 1997; Woloshuk and Prieto, 1998) and thirteen of these have been assigned enzymatic roles.

In contrast to the conservation in the gene order and distances between the *A. parasiticus* and *A. flavus* AF clusters, the ST cluster shows no such conservation. At present twelve *A. nidulans stc* genes show significant identity to genes present in the AFB_1 cluster in *A. parasiticus* and *A. flavus* (Figure 1.3; Table 1 in Brown *et al.*, 1996). Unlike the ST cluster, the AFB_1 cluster has not been fully sequenced. Completion of this sequence will permit detailed comparison between the two clusters and will also determine whether genes involved in other types of AF biosynthesis (AFB_2, AFG_x etc.) are within this cluster.

Some strains of *A. parasiticus* contain a 12 kb duplicated chromosomal region containing *ver1* and *aflR* (Liang *et al.*, 1996; Klich *et al.*, 1997). This second *ver* gene (*ver1B*) displays 95% amino acid identity with *ver1*, but contains a premature translational stop codon (residue 87) and encodes a non-functional truncated polypeptide (Liang *et al.*, 1996). This duplication has not been found in any *A. flavus* and *A. nidulans* strains to date.

1.6.3 Expression and Regulation of the ST/AF Pathway Genes.

Expression of the genes in the ST/AF biosynthetic cluster is regulated by the *aflR* gene product. The *aflR* gene was cloned in *A. flavus* by complementation analysis (initially termed *afl-2*; Payne *et al.*, 1993) and in *A. parasiticus* on the basis of overproduction of pathway intermediates following transformation of the fungus with a cosmid that contained an extra copy of the gene (initially termed apa-2; Chang *et al.*, 1993). These two genes share 96% nucleotide identity (97% amino acid identity) and are functional homologs (Chang *et al.*, 1993, Woloshuk *et al.*, 1994). Southern blot analysis indicated that a copy of the *aflR* gene was present in the closely related *A. oryzae* and *A. sojae* genomes (Woloshuk *et al.*, 1994).

An *aflR* homolog is present in the *A. nidulans* ST biosynthetic gene cluster, but in contrast to the high identity between the *A. flavus* and *A. parasiticus* genes, the *A. nidulans aflR* shares only 33% amino acid sequence identity, although there is 71%

amino acid identity over the zinc binuclear cluster region. Interestingly, the *A. flavus aflR* was able to activate ST expression in *A. nidulans* (Yu *et al.*, 1996).

The *aflR* gene encodes a 47 KDa DNA-binding protein, AFLR. This contains a Gal4-type binuclear zinc finger cluster with the Cys-$(Xaa)_2$-Cys-$(Xaa)_6$-Cys-$(Xaa)_6$-Cys-$(Xaa)_2$-Cys-$(Xaa)_6$-Cys motif which confers DNA binding specificity. AFLR also contains a highly acidic domain adjacent to the carboxyl terminus, similar to GAL4, which is thought to be important in transcriptional activation of other pathway genes. *aflR* may be autoregulated as AFLR has been shown to bind specifically to a palindromic sequence (TTAGGCCTAA) 120 bp upstream of the *aflR* translation start site (Chang *et al.*, 1995). However this sequence is not present in all pathway promoters and there are no reports of AFLR binding to other biosynthetic gene promoters.

AFLR is absolutely required for expression of the ST/AFB_1 biosynthetic genes and ST/AF production (Payne *et al.*, 1993; Chang *et al.*, 1993; Woloshuk *et al.*, 1994; Yu *et al.*, 1996). This was also well demonstrated by Prieto *et al.* (1996) who complementated the *A. flavus* mutant *afl-1* (which contains a 120 kb deletion spanning the whole AFB_1 biosynthetic cluster) with three overlapping cosmid clones, only one of which (clone 20B11) contained a copy of *aflR*. Only transformants containing the 20B11 clone could express pathway genes and produce AFB_1. This also illustrated that clustering of pathway genes is not required for efficient AF production. From a different angle, studies in which *aflR* was overexpressed by placing it under the control of different promoters showed that ST/AFB_1 could be synthesised under conditions which usually do not support ST/AFB_1 production (Yu *et al.*, 1996; Chang *et al.*, 1995; Flaherty and Payne, 1997; Feng and Leonard, 1998).

In culture, *A. parasiticus* and *A. flavus* produce AFs during idiophase, the transition from exponential growth phase to stationary phase (characteristic of secondary metabolites). Many of the enzymatic activities required for AFB_1 biosynthesis have been identified specifically from this stage (Cleveland *et al.*, 1987; Lin and Anderson, 1992; Skory *et al.*, 1993; Anderson and Green, 1994; Silva and Townsend, 1996; Matsushima *et al.*, 1994). AF biosynthetic genes follow the same temporal pattern in a coordinated manner (Brown

et al., 1996; Skory *et al.*, 1993; Liang *et al.*, 1997; Klich *et al.*, 1997; Flaherty *et al.*, 1995; Yu *et al.*, 1996).

Flaherty and Payne (1997) investigated what effect altering the temporal pattern of *aflR* expression would have on pathway gene expression and AF production. The *aflR* gene was fused to the *A. nidulans* glyceraldehyde-3-phosphate dehydrogenase (*gpdA*) gene promoter to produce the constitutively expressed *gpdA*(P)::*aflR* construct. When transformed into the *A. flavus* 656-2 mutant (Payne *et al.*, 1993) which contains a mutated copy of *aflR* and does not produce AF, constitutive expression of biosynthetic genes was observed. A similar experiment was performed in *A. flavus* 86-10 which contains a functional copy of *aflR* and produces wild type levels of AF. Once again, constitutive expression of the genes was observed. However in both cases the temporal profile of AF accumulation did not change. This implies that transcriptional regulation of the biosynthetic pathway by *aflR* is not the only mechanism involved in the control of AF production (Flaherty and Payne, 1997).

Feng and Leonard (1998) investigated what effect various culture conditions had in ST/AFB_1 biosynthesis in *A. nidulans* and *A. parasiticus*. An increase in the culture temperature reduced both AF production and pathway gene expression in *A. parasiticus* but the opposite was true in *A. nidulans*. The same pattern was observed with the nitrogen source. With ammonium as the sole nitrogen source, *A. nidulans* did not produce ST (mycelial growth stopped after three days), but AFB_1 production and gene expression was detected in *A. parasiticus*. These results further support the theory that transcriptional regulation by *aflR* is not the only mechanism involved in the regulation of ST/AF biosynthesis, and that these factors could differ between *A. nidulans* and *A. parasiticus*.

Guzmán-de-Peña and Ruiz-Herrera (1997) established a correlation between AF biosynthesis and sporulation in *A. parasiticus*. Mutants blocked in sporulation were unable to produce AF. When this block was relieved and sporulation induced, so too was AF production. Hicks *et al.* (1997) presented a model in which both asexual sporulation and mycotoxin production require inactivation of proliferative growth through inhibition of FadA-dependent signalling in *A. nidulans*. *FadA* encodes the α subunit of a

heterotrimeric G protein. The *Magnaporthe grisea magB* gene product displays 93% amino acid identity to that of *fadA* and also encodes a heterotrimeric G protein α subunit (Liu and Dean, 1997). Disruption of *magB* in M. grisea inhibited appresorium formation, blocked sexual development and significantly reduced vegetative growth and pathogenicity, implying that G protein α subunits are involved in signal transduction pathways controlling these processes (Liu and Dean, 1997).

Other factors affecting AF biosynthesis have also been reported. Flaherty *et al.* (1995) found that a heat-stable 10 KDa molecule derived from maize (possibly a complex seed molecule degraded by fungal extracellular enzymes) was an inducer of AF biosynthesis in *A. flavus*. The exact identity of this molecule is not known. Similarly, seed lipooxygenase products, including 13*S*-hydroperoxy fatty acids, directly or indirectly repressed ST/AF biosynthesis in *A. nidulans* and *A. parasiticus* (Burow *et al.* 1997).

From these reports it is apparent that *aflR* is critical for ST/AFB_1 biosynthesis. However, other factors including sporulation, growth, environmental conditions and possibly plant signal molecules are also important for ST/AF production. Whether particular physiological conditions are required specifically for the biosynthesis of ST/AF or globally for secondary metabolism remains to be determined. Because ST/AF is not required for the growth of the fungus, it is fascinating that such seemingly complex regulation exists for these dispensable pathways. It will be very interesting for future analysis to investigate such regulation in the dothistromin system.

Recent research in filamentous fungi has identified that genes involved in pathways that are either not required for growth, like ST/AFB_1 biosynthesis, or are only required for growth only under a limited range of conditions, are often clustered. Such pathways have been termed "dispensable" (Keller and Hohn, 1997) or "weakly selected" (Lawrence and Roth, 1996). This general observation, along with the seemingly close relationship between dothistromin and AF biosynthesis, indicates that dothistromin biosynthetic genes are also clustered and if so, what is the evolutionary relationship between the two pathways? Gene clusters and possible mechanisms of evolution and transfer are discussed below.

1.7 GENE CLUSTERS.

Gene clusters are defined as the close linkage of two or more genes that participate in a common metabolic or developmental pathway. Gene clusters are prominent features of bacterial chromosomes but are rare in eukaryotic systems. However, as mentioned above, recent research has identified a number of examples of gene clusters in filamentous fungi (reviewed in Keller and Hohn, 1997).

Penicillins and cephalosporins are β-lactam-containing antibiotics produced by fungi and bacteria. Penicillin biosynthetic genes are present within a 20 kb cluster in *Penicillium chrysogenum*, *A. nidulans* and the cephalosporin-producing *Acremonium chrysogenum*. It is thought that at least part of this pathway may have been acquired through a horizontal gene transfer event from a prokaryote. The isopencillin N (IPN) synthetase genes of the three fungi display 80% identity with each other and 60% identity to prokaryotic IPN synthetase genes. The fungal IPN synthetase genes lack introns and have a high GC content, which is indicative of a bacterial origin (Weigel *et al.*, 1988) and phylogenetic analysis supports this theory (Section 1.7.2; Buades and Moya, 1996).

Melanin is a polyketide-derived dark brown or black pigment produced by a wide range of fungi which accumulates in fungal cell walls. Melanin is considered to confer tolerance to environmental stresses (such as UV radiation), and mediates the forces for mechanical penetration of appressoria (Perpetua *et al.*, 1996). Melanin biosynthetic genes are linked in *Alternaria alternata* (Kimura and Tsuge, 1993), unlinked in *Colletotrichum lagenarium* (Kubo *et al.*, 1996) and *Magnaporthe grisea* (Chumley and Valent, 1990) and partially linked in *Cochliobolus miyabeanus* and *Cochliobolus heterostrophus* (Kubo *et al.*, 1996; but conflicted by Shimizu *et al.*, 1997). This illustrates that gene clustering is not required for efficient secondary metabolite production.

Other examples of gene clusters in filamentous fungi are; the Quinate/shikimate pathway in *Neurospora crassa* and *A. nidulans*, ethanol utilisation in *A. nidulans*, proline utilisation in *S. cerevisiae* and *A. nidulans*, nitrate assimilation in *N. crassa* and *A. nidulans*, trichothecene and macrocyclic trichothecene production in *Fusarium sporotrichoides* and *Myrothecium roridum* respectively (reviewed in Keller and Hohn, 1997).

1.7.1 Evolution of Gene Clusters.

Four general models have been proposed for the origin of gene clusters clustered (Lawrence and Roth, 1996; Lawrence, 1997). (1) The Natal model proposes that gene clusters originate *in situ* by gene duplication and divergence, therefore gene position and clustering is a historical property and provides no direct benefit to the individual. An example of this is the mammalian β-globin gene cluster. (2) The Fisher model proposes that genes will cluster when the two genes work well together (co-adaptation). Evidence for this is that genes in bacteriophage genomes are within logical groups, for example genes for head proteins are clustered. (3) The co-regulation model proposes that gene clusters facilitate coordinate expression and regulation which provides a selective benefit to the individual. Bacterial operons are examples of this. (4) The selfish operon model proposes that gene clusters allow dissemination of functionally related genes via horizontal transfer. Gene clusters do not provide a selective advantage to the organism, but the physical proximity provides a strong advantage to the genes themselves when competing via lateral transmission with unclustered alternative alleles ("selfish" behaviour). This clustering is advantageous only to the genes themselves, not to the immediate host organism, which is consistent with the observation that predominantly "disposable" pathway genes are clustered (Lawrence and Roth, 1996; Lawrence, 1997).

Prade *et al.* (1997) proposed horizontal transfer of selfish gene clusters (including the ST cluster), and dispersal by transposition, as an explanation for the nonrandom pattern of repeats along the eight chromosomes of *A. nidulans* observed through the reconstruction of the physical map. *Stc* genes are embedded in a repetitive DNA region on chromosome four which is devoid of essential genes, and the same pattern is seen for other clusters in *A. nidulans*. This is expected from cluster colonisation.

1.7.2 Horizontal Gene Transfer.

Horizontal or lateral gene transfer is defined as the transfer of genetic material from one genome to another, specifically between species. The most studied and well known example of horizontal gene transfer is the transmission of T-DNA from the bacterium *Agrobacterium tumefaciens* to a plant cell genome. Horizontal transfer is common in bacteria and is proposed to be a major influence in bacterial evolution (Lawrence and Roth, 1996; Lan and Reeves, 1996). An increasing number of examples of horizontal

transfer events are being proposed for eukaryotic genes. An example of horizontal transfer involving filamentous fungi is the proposed transfer of the IPN synthetase gene from bacteria to fungi (Buades and Moya, 1996).

The mechanism(s) by which horizontal transfer is thought to have been achieved is very difficult to detect as many signals have been lost or diminished over time. Transfer to eukaryotes from prokaryotes could be mediated by conjugation (as in T-DNA transfer) or symbiosis. Transfer between eukaryotes could be mediated by retroviruses, transposons (including retrotransposons, retroposons) or parasites as vectors carrying mobile elements (as in *Drosophila* species, where mites were thought to be the vector for P-element transfer; Li, 1997).

Detection of horizontal transfer events is difficult, and even harder to prove. Usually it is proposed following the discovery of an outstanding discontinuity; for example a vastly different GC content, lack of introns and/or no counterpart in other fungi. There are three main methods to test for horizontal transfer events, of which two are complex statistical methods and one is based on phylogenetics (discussed in Li, 1997). In the latter method, two phylogenetic trees are constructed. One is based on sequences representative of a species (for example rDNA sequences) which give rise to a "species tree" or "true tree". The other is a "gene tree" or "inferred tree" constructed from the sequence under investigation. If the two trees are incongruous and the support for the inferred tree is greater than the true tree, this is supportive of a horizontal transfer event having occurred. This analysis has been used positively in the IPN synthetase transfer investigation as well as other examples involving filamentous fungi (Buades and Moya, 1996; Masel *et al.*, 1996; Hibbett, 1996).

1.8 ISOLATION OF A PUTATIVE DOTHISTROMIN GENE BASED ON SEQUENCE SIMILARITY.

The conversion of VA to the toxic ST is a critical step in the production of ST/AFB_1. The gene involved in this conversion, *ver1*, has undergone extensive study. Genetic disruption of *ver1* resulted in the accumulation of VA and no production of ST (Skory *et al.*, 1992). However, trace amounts of AFB_1 are still observed indicating the possibility of alternative routes for AF biosynthesis (Skory *et al.*, 1992; Keller *et al.*, 1994). *ver1* encodes a

putative ketoreductase and displays significant amino acid identity to ketoreductase genes involved in melanin biosynthesis, and is highly conserved between *A. parasiticus*, *A. flavus* and *A. nidulans* (Brown *et al.*, 1996).

These characteristics made *ver1* an attractive candidate gene for isolation in *D. pini*. To determine whether *D. pini* contained a *ver*-like gene, a Southern blot of *D. pini* genomic DNA was hybridised with a labelled 1.8 kb *ver1* gene fragment from *A. parasiticus*. The results indicated that *D. pini* did contain a *ver*-like gene. This heterologous probe was then used to screen a *D. pini* genomic library and a lambda clone, λCGV1, was isolated. Preliminary sequencing confirmed the presence of a *ver*-like gene on λ CGV1 (Gillman, 1996).

A transformation system for *D. pini* has been designed (Bidlake, 1996). This will allow the targeted gene disruption of genes such as the *ver*-like gene in *D. pini* to confirm the function of the gene and to create dothistromin-minus mutants.

1.9 AIMS AND OBJECTIVES

The primary focus of current *Dothistroma* research is to obtain mutants blocked in the synthesis of dothistromin via targeted gene disruption of biosynthetic genes. This will allow direct assessment of the role of dothistromin in pathogenicity of *D. pini*. It will also provide a detailed molecular picture of toxin (dothistromin) biosynthesis and regulation, and will provide insights on plant-toxin interactions. An interesting focal point upon further characterisation will be the comparison between dothistromin and AF biosynthesis.

This study focuses on the characterisation of genes involved in dothistromin biosynthesis. This will be achieved by complete sequencing of λCGV1 which has been shown to contain a *ver*-like sequence (Section 1.8; Gillman, 1996). This will indicate whether likely candidate genes for dothistromin biosynthesis are clustered and will allow preliminary comparison to other systems. Expression of genes identified will be investigated to determine whether a coordinated expression pattern is observed. Large scale sequencing will also allow analysis into the evolution of the pathway to be performed.

2. MATERIALS AND METHODS.

2.1 FUNGAL AND BACTERIAL STRAINS, LAMBDA CLONES AND PLASMIDS.

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

2.2 MEDIA.

All media were prepared with Milli-Q water and sterilised by autoclaving at 121°C for 15 minutes. Liquid media were cooled to room temperature before addition of antibiotic(s) and inoculation. Solid media were cooled to approximately 50°C before addition of supplements and pouring.

2.2.1 Luria Broth.

Luria Broth (LB) media contained (g/L): tryptone, 10; NaCl, 5; yeast extract (Oxoid), 5. The pH was adjusted to 7.4 before autoclaving. For solid media, agar was added to 15 g/L. When required, ampicillin, isopropylthio-β-D-galactoside (IPTG) and 5-bromo 4-chloro 3-indolyl-β-D-galactoside (X-gal) were supplemented at final concentrations of 100, 30 and 60 μg/ml respectively.

2.2.2 Terrific Broth.

Terrific Broth (TB) was prepared by adding 47g TB (Life Technologies) to 1 L water.

2.2.3 *Dothistroma* Media (DM).

Contained (g/L): nutrient agar (Oxoid), 23; and malt extract (Oxoid), 50.

2.2.4 Dothistroma Broth (DB).

Contained (g/L): nutrient broth (Oxoid), 23; and malt extract (Oxoid), 50.

2.2.5 Dothistroma Sporulation Media (DSM).

Contained (g/L): malt extract (Oxoid), 15; yeast extract (Oxoid), 5; and agar, 20.

Table 2.1 Strains, λ clones and plasmids.

Strain, Plasmid, λ Clone.	Relevant Characteristics	Source or Reference.
Fungal Strain.		
Dothistroma pini.		
Dp2	Single spore isolate, basic laboratory strain.	Long Mile Road, NZFRI, Rotorua
Bacterial Strains.		
Escherichia coli.		
XL1-Blue	*sup*E44 *hsd*R17 *rec*A1 *end*A1 *gyr*A46 *thi* *rel*A1 *lac*⁻ F` [*pro*AB⁺ *lac*I^{q}Δ (*lac*Z) M15 Tn10(*tet*^{r})]	Bullock *et al.*, 1987.
DH-1	F` *rec*A1 *end*A1 *gyr*A96 *thi*-1 *hsd*R17 (r⁻m⁺) *sup*E44 *rel*A1 λ-	Hanahan, 1983.
λ clones		
λBT-1	LambdaGEM®-12 (Promega) containing genomic DNA (including *tub-1*) from Dp2.	Bidlake, 1996.
λCGV1	LambdaGEM®-12 (Promega) containing genomic DNA (including *ver1*) from Dp2	Gillman, 1996.
Plasmids.		
pUC118	Amp^r *lac*Z` (3.2 kb)	Messing, 1983.
pUC18	Amp^r *lac*Z` (2.7 kb)	Yanish-Perron *et. al.*, 1995.
pGEM®-T	Amp^r *lac*Z` (3.0 kb)	Promega.
R130	pUC118 containing a 0.75 kb *Sal*I fragment from λCGV1.	Gillman, 1996.
R134	pUC118 containing a 0.75 kb *Eco*RI fragment from λBT-1.	This study.
R135	pUC118 containing a 0.97 kb *Eco*RI fragment from λBT-1.	This study.
R136	pUC118 containing a 1.1 kb *Xho*I fragment from λBT-1.	This study.
R137	pUC118 containing a 1.45 kb *Eco*RI fragment from λBT-1.	This study.
R138	pUC118 containing a 2.7 kb *Xho*I fragment from λBT-1.	This study.
R139	pUC118 containing a 0.4 kb *Xho*I fragment from λBT-1.	This study.
R140	pUC118 containing a 5.8 kb *Eco*RI-*Bam*HI fragment from λBT-1.	This study.
R149	pUC18 containing a 0.6 kb *Eco*RI-*Bam*HI	This study.

	fragment from λCGV1.	
R150	pUC18 containing a 0.9 kb *Eco*RI-*Bam*HI fragment from λCGV1.	This study.
R151	pUC18 containing a 2.1 kb *Eco*RI fragment and a 0.25 kb *Eco*RI-*Bam*HI fragment from λCGV1.	This study.
R152	pUC18 containing a 3.2 kb *Eco*RI-*Bam*HI fragment from λCGV1.	This study.
R153	pUC18 containing a 1.1 kb *Eco*RI-*Xho*I fragment from λCGV1.	This study.
pMF4	pGEM®-T containing a 4.0 kb fragment amplified from λCGV1 using primers DS152ep1 and 151Rep3	This study.
pMF800	pGEM®-T containing a 0.8 kb fragment amplified from λCGV1 using primers SP6 and 151Fep5	This study.

2.3 GROWTH AND MAINTENANCE OF CULTURES.

E. coli cultures were grown at 37°C overnight on LB agar plates, or in LB Broth, with the appropriate selection supplements (Section 2.2.1). Plates were sealed with parafilm and stored at 4°C. All cultures were regularly subcultured by streaking onto fresh LB plates. For long-term storage, glycerol stocks of selected *E. coli* cultures were prepared by pelleting cells from an overnight liquid culture by a 1 min centrifugation (in a microcentrifuge), then resuspended in 1 ml of 20 % (v/v) glycerol and stored at -70°C.

D. pini cultures were grown on DM or DSM plates at 20°C in the dark for 10-14 days. For liquid cultures, DB was inoculated with 10^5-10^6 spores per ml, and incubated with shaking (150 rpm) for 7-14 days.

2.4 COMMON BUFFERS AND SOLUTIONS.

All solutions were prepared with Milli-Q water and sterilised by autoclaving at 121°C for 15 minutes, unless otherwise stated.

2.4.1 TE Buffer.

Contained 10 mM Tris-HCl and 1 mM Na_2EDTA (TE 10:1) or 10 mM Tris-HCl and 0.1 mM Na_2EDTA (TE 10:0.1); and was prepared to the required concentration from 1 M Tris-HCl (pH 8.0) and 250 mM Na_2EDTA (pH 8.0) stock solutions.

2.4.2 1 x TAE Buffer.

Contained 40 mM Tris-HCl, 2 mM Na_2EDTA and 20 mM acetic acid.

2.4.3 1 x TBE Buffer.

Contained 89 mM Tris-HCl, 2.5 mM Na_2EDTA and 89 mM boric acid, pH 8.3

2.4.4 10 x Sequencing TBE Buffer.

Contained (g/L): Tris, 162; Na_2EDTA, 9.5; and boric acid, 27.5.

2.4.5 20 x SSC.

Contained 3 M NaCl and 0.2 M tri-sodium citrate. 2 x SSC was prepared by diluting 20 x SSC with Milli-Q water.

2.4.6 Ethidium Bromide.

The ethidium bromide solution used for staining of agarose gels was prepared by adding 1 µl of a 10 mg/ml stock per 10 ml of Milli-Q water to give a final concentration of 1 µg/ml.

2.4.7 RNase A (DNase free).

Contained 10 mg/ml RNase A (Sigma) in 10 mM Tris-HCl (pH 7.5), 15 mM NaCl. The solution was heated to 100°C for 15 min, then allowed to cool slowly before being dispensed into aliquots and stored at -20°C.

2.4.8 STET Buffer.

Contained 8% (w/v) sucrose, 5% (v/v) Triton X-100, 50 mM Na_2EDTA (pH 8.0) and 50 mM Tris-HCl (pH 8.0).

2.4.9 Gel Loading Buffer (10 x).

Contained 2 M Urea, 50 % (v/v) glycerol, 50 mM Tris acetate, 0.4 % (w/v) Bromophenol Blue and 0.4 % (w/v) Xylene cyanol.

2.4.10 Acrylamide mix

Contained (g/L): urea, 480; acrylamide, 57; bis-acrylamide, 3. The mix was made up to around 800 ml and deionised with Amberlite MB-3 (Sigma), then filtered through a sintered glass funnel (porosity 1). 100 ml of 10 x sequencing TBE buffer (Section 2.4.4) was then added and the volume made up to 1 litre with Milli-Q water.

2.4.11 Hybridisation solution.

Contained (per litre): 150 ml of 20 x SSC (Section 2.4.5), 20 ml of 50 x Denhardt's solution (Section 2.4.12), 30 ml of 10 % (w/v) SDS and 5 ml of salmon sperm DNA (at 50 µg/ml).

2.4.12 50x Denhardt's Solution.

Contained (g/L): ficoll, 10.0; polyvinylpyrolidone, 10.0; bovine serum albumin, 10.0. Filter sterilised.

2.4.13 10 X MOPS Buffer.

Contained (g/L): MOPS (3-[N-morpholino] propanesulphonic acid), 41.2; sodium acetate, 3-hydrate, 10.9; Na_2EDTA, 3.7. The pH was adjusted to 7.0 with NaOH (prepared in nuclease-free distilled water). Stored at room temperature, protected from light.

2.5 DNA ISOLATION.

2.5.1 Rapid Boiling Plasmid Preparation.

This method for preparing plasmid DNA was performed as described by Holmes and Quigley (1981). The DNA was resuspended after drying in 25µl sterile Milli-Q water or TE (10:0.1; Section 2.4.1).

2.5.2 Alkaline Lysis Plasmid Preparation (Small Scale).

This method for preparing plasmid DNA was performed as described by Sambrook *et al.* (1989). The DNA was resuspended after drying in 25µl sterile Milli-Q water or TE (10:0.1; Section 2.4.1). This method produced plasmid DNA of higher quality than the rapid boiling method and was used to isolate plasmid DNA that was to be manually sequenced.

2.5.3 Modified Alkaline-Lysis, PEG Precipitation Procedure (Small-Medium Scale).

This procedure was used for preparation of plasmid DNA that was to be automatically sequenced (Section 2.14.2). This method was performed according to instructions provided with the *Taq* DyeDeoxy Terminator Cycle Sequencing Kit (Perkin Elmer).

2.5.4 DNA Isolation from Fungal Cultures (Large Scale).

Isolation of genomic DNA from *D. pini* cultures was based on the method of Byrd *et al.*, (1990). In a pre-cooled mortar and pestle 100 mg of freeze dried mycelium was ground to

a fine powder under liquid nitrogen, then suspended in 10 ml of extraction buffer (150 mM Na_2EDTA, 50 mM Tris-HCl pH 8.0, 1% [w/v] sodium lauroyl sarcosine and 2 mg/ml proteinase K). The solution was centrifuged at 2000 **g** for 10 min (all centrifugations were at 4°C) and the supernatant incubated at 37°C for 20 min. Two volumes of phenol were added, mixed and centrifuged at 2500 **g** for 15 min. The aqueous phase was transferred to a clean tube and the process repeated this time with one volume of phenol/chloroform, then again with 2 volumes of chloroform. The aqueous phase was then centrifuged at 2500 **g** for 20 min to remove any polysaccharides. DNA was precipitated by adding 1 volume of isopropanol to the supernatant, incubated on ice for 30 min then centrifuged at 1600 **g** for 10 min. DNA was washed with 70% ethanol, dried, resuspended in 100 µl of Milli-Q water and quantified (Section 2.7).

2.6 PURIFICATION OF DNA.

2.6.1 Phenol/Chloroform Extraction.

DNA samples were extracted with equal volumes of Tris-equilibrated phenol (United States Biochemical) and chloroform, mixed thoroughly and centrifuged (≥ 15000 **g**). The aqueous phase was re-extracted until a clear interface between the two phases was obtained. Samples were then extracted once with two volumes of chloroform. DNA was then isolated by ethanol precipitation (Section 2.6.2)

2.6.2 Ethanol Precipitation.

To concentrate DNA solutions, one tenth volume of 3 M sodium acetate and either 2.5 volumes of 95% ethanol, or 0.6 volumes of isopropanol, were added, mixed and left on ice for at least 15 min. The DNA was then pelleted by centrifugation at 15000 **g** for 10 min. The pellet was washed once with 70% ethanol, then dried under vacuum. The DNA was resuspended in Milli-Q water or TE (10:0.1; Section 2.4.1).

2.6.3 Agarose Gel Purification of DNA Fragments.

DNA (usually a restriction enzyme digest or a PCR reaction) containing the fragment to be purified was run on a 1-1.5% SeaPlaque (FMC) low melting point agarose gel in TAE buffer (Sections 2.9 and 2.4.2). After staining in ethidium bromide, the DNA was visualised under long wave UV light, and the appropriate band excised using a clean

scalpel blade. The DNA was extracted from the agarose using the QIAquick Gel Extraction Kit (Qiagen) according to the manufacturers instructions.

2.6.4 Purification of DNA from a PCR Reaction.

DNA was purified from a PCR reaction either by; gel purification (Section 2.6.3), or the QIAquick PCR Purification Kit (Qiagen) according to manufacturers instructions.

2.7 DETERMINATION OF DNA CONCENTRATION

DNA was quantified by one or more of the following methods.

2.7.1 Determination of DNA concentration by spectrophotometric assay.

Spectrophotometric quantification was used mostly for pure DNA samples of high concentration. Each sample was diluted appropriately and the absorbance measured at both 260 nm and 280 nm. The DNA concentration was calculated on the assumption that an absorbance of 1.0 at 260 nm is equivalent to 50 µg/ml. Purity was indicated by the 260 nm/280 nm ratio, where pure DNA has a ratio of 1.8.

2.7.2 Determination of DNA concentration by fluorometric assay.

Fluorometric quantification was used for pure or impure DNA samples of low to high concentrations. DNA was quantified on a Hoefer Scientific TKO 100 Fluorometer according to the manufacturer's protocol. The scale of the fluorometer was set to 100 using 2 µl of 100 µg/ml calf thymus DNA added to 2 ml of a dye solution containing 1 x TNE buffer (10 mM Tris-HCl, 1 mM Na_2EDTA and 100 mM NaCl, pH 7.4) and 0.1 µg/ml Hoechst 33258 dye. Once the scale was reliably set, 2µl of sample DNA was added to 2 ml of the dye solution, and the resulting value recorded as the concentration of DNA in ng/µl (if there was enough DNA sample this assay was performed in triplicate).

2.7.3 Determination of DNA concentration by Gel Electrophoresis.

A series of Lambda DNA or pUC118 DNA concentration standards were run on an agarose gel (Section 2.9) alongside the DNA sample. The DNA sample concentration was estimated by comparing the intensity of ethidium bromide fluorescence to that of the DNA concentration standards. This method was performed for most samples as it

confirmed the presence of DNA and suggested which assay would be better if more accurate quantification was required.

2.8 RESTRICTION ENDONUCLEASE DIGESTION OF DNA.

Restriction endonuclease digests were performed at 37°C, in the buffer specified by the manufacturer, for a minimum of 1 hour (and maximum of overnight). Digestion was performed with 1-10 units of restriction enzyme per µg DNA. The enzyme volume never exceeded one tenth of the total reaction volume. When necessary, RNA was removed from DNA samples by the addition of RNaseA (Section 2.4.7) to a final concentration of 1.0 µg/ml and incubated at 37°C for 10-30 min. Digests were checked by running an aliquot on an agarose gel (Section 2.9). If digestion was not complete, further restriction enzyme was added. Digests were stored at -20°C.

2.9 AGAROSE GEL ELECTROPHORESIS OF DNA.

DNA fragments were size fractionated by electrophoresis through 0.75%-2.5% agarose dissolved in 1 x TAE (Section 2.4.2) or TBE (Section 2.4.3) buffer at 75-100 volts. Large agarose gels were usually run at 25 volts overnight. Dyes present in the gel loading buffer (Section 2.4.9), which was added to DNA samples before loading, allowed estimation of DNA migration. After electrophoresis, agarose gels were stained in ethidium bromide for 10-30 minutes (depending on the size of the gel), washed in water, observed under short wave UV light and photographed.

2.10 DETERMINATION OF DNA MOLECULAR WEIGHTS.

DNA fragments were sized by running the DNA sample on an agarose gel alongside known size ladders such as; *Hind*III/*Eco*RI double digest of lambda DNA, 100 bp ladder (Life Technologies), 1 kb ladder (Life Technologies). The mobility of the standard markers from the wells were measured and used to determine the molecular weight of the unknown fragments by comparing the relative mobility between the lanes.

2.11 DNA SUB-CLONING TECHNIQUES.

Fragments to be sub-cloned were purified (Section 2.6), usually by gel purification, then ligated into the appropriate vector (Section 2.11.1), and transformed into a suitable *E. coli* host strain (Section 2.11.2,). Blue/white colour selection was used for initial

screening of transformants. When numbers of suspected transformants were large, PCR (Section 2.13) using appropriate primers was performed to further screen transformants. The presence of insert was always confirmed by isolation of plasmid DNA (Sections 2.5.1-2.5.3) and restriction enzyme digestion (Section 2.8).

To increase cloning efficiency, directional cloning was used when possible and the vector was gel purified after restriction enzyme digestion.

2.11.1 Ligation Reactions.

Ligation reactions were performed in a total reaction volume of 20 µl, containing 2 µl of 10 x ligase buffer, 0.5 µl of T4 DNA Ligase (New England Biolabs) and a 2:1 molar ratio of insert:vector (usually with a minimum of 20 ng of vector). Ligation reactions were either placed at 4°C overnight, 14°C for 3 hours, or at room temperature for a minimum of 1 hour. Ligation reactions were checked for completion by removing a 1 µl aliquot and comparing this on an agarose gel with an aliquot taken before addition of ligase. The presence of higher molecular weight bands in the ligase sample, as well as a reduction in intensity of the insert DNA band, were characteristics of a successful ligation.

2.11.2 Transformation of *E. coli*.

2.11.2.1 Electroporation of *E. coli*.

2.11.2.1.1 Preparation of Electroporation Competent *E. coli* Cells.

One litre of LB broth was inoculated with 10 ml of an overnight culture of the desired *E. coli* strain (XL-1 or DH1), and grown at 37°C shaking to mid-log phase (OD_{600} 0.5-1.0). The cells were chilled on ice for 20 min then harvested by centrifugation for 10 min at 4000 **g** (all centrifugation steps performed at 4°C). The cells were washed sequentially (by resuspension, centrifugation at 4000 **g** to pellet the cells and removal of the supernatant) in 1.0 litre and 0.5 litre of ice cold sterile water, 20 ml of ice cold 10% (v/v) glycerol, then finally resuspended in 4 ml of ice cold 10% (v/v) glycerol. Cells were stored at -70°C in 200 µl aliquots.

2.11.2.1.2 Electroporation.

4 µl of DNA (ligation mixture or controls) was added to 40 µl of electroporation competent *E. coli* cells (Section 2.11.2.1.1) in an ice cold eppendorf tube and gently mixed. This was then immediately transferred to an ice cold 0.2 cm electroporation cuvette, tapped to the bottom and electroporated in a Biorad gene pulsar set at 25µF, 2.5 kV and 200 Ω. 220 µl of LB broth (Section 2.2.1) was immediately added to the cells, mixed and transferred to fresh tubes which were incubated at 37°C for 1 hour. Positive (10 ng circular plasmid DNA) and negative (water only) controls were always included. Cells were spread onto LB agar plates containing appropriate supplements (Section 2.2.1) and incubated overnight at 37°C.

2.11.2.2 Calcium Chloride Transformation of *E. coli* (Heat Shock Method).

Preparation of calcium chloride competent cells and transformation of plasmid DNA was performed according to Cohen *et al.* (1972) and Ausubel *et al.* (1994). Transformed cells were grown at 37°C overnight on LB plates containing the appropriate supplements (Sections 2.2.1 and 2.3).

2.12 SOUTHERN BLOTTING AND HYBRIDISATION.

2.12.1 Southern (Capillary) Blotting.

DNA to be blotted and hybridised was electrophoresed through a large 1-1.5 % agarose gel, stained in ethidium bromide and photographed (with a ruler next to the gel to allow molecular weight size determination). The gel was then gently agitated in depurination solution (250 mM HCl) for 15 min, then denaturation solution (500 mM NaOH, 500 mM NaCl) for 30 min, and neutralisation solution (500 mM Tris [pH 7.4], 2 M NaCl) for 2 x 20 min. Finally, the gel was washed in 2 x SSC (Section 2.4.5) for 3 min before placing on the blotting apparatus as described by Southern (1975) and Ausubel *et al.* (1994).

After blotting overnight, the apparatus was disassembled and the membrane (Hybond-N, Amersham) left to air dry between 3MM paper, then the DNA was fixed to the membrane by baking under vacuum at 80°C for two hours. Finally excess salt was

removed from the membrane by washing in 2 x SSC for 5 min. The membrane was stored at 4°C.

2.12.2 Preparation of [α-^{32}P]dCTP-Labelled DNA Probe.

Probe DNA (25 ng) was labelled with [α-^{32}P]dCTP (Amersham) using the Ready-To-Go DNA Labelling Kit (Pharmacia) according to the manufacturers instructions. Unincorporated nucleotides were removed by using the ProbeQuant G-50 Micro Columns (Pharmacia) as described by the manufacturers.

2.12.3 Southern Blot Hybridisation.

The Southern blot to be probed was prehybridised for 2 hours at 65°C in a sealed glass hybridisation tube containing approximately 7 ml of hybridisation solution (Section 2.4.11). After prehybridisation, the hybridisation solution was replaced with fresh solution (prewarmed to 65°C), and the denatured labelled probe was carefully added. Hybridisation was performed in a rotating hybridisation oven at 65°C overnight.

After hybridisation, the solution was replaced with 30-50 ml of wash solution (0.1-2.0 x SSC, 0.1 % SDS). The membrane was washed 3 times for 25 minutes at 65°C, then wrapped in gladwrap. This was then placed in a X-ray cassette with intensifying screens against either slow film (Fuji Medical) or fast film (Kodak Scientific Imaging) X-ray film. After exposure for an appropriate period of time at -70°C, the film was developed in a dark room by placing in developing solution for 5 min and fixing solution for at least 3 min, then rinsed in water and dried.

2.12.4 Removal of [α-^{32}P]dCTP-Labelled DNA from Southern Blots (Stripping).

Membranes to be stripped were placed in a container containing 0.1% SDS at 100°C and left shaking for 30 min (repeated twice). Stripping was confirmed by autoradiography as described in Section 2.12.3. If stripping was incomplete, the process was repeated.

2.13 AMPLIFICATION OF DNA BY THE POLYMERASE CHAIN REACTION (PCR).

PCR reactions were set up on ice using a cocktail which contained all common reagents used for n+1 PCR reactions. Uncommon reagents were pipetted separately. PCR

reactions were in a total volume of 25 µl in 0.2 ml PCR tubes. The final concentrations of each component in 1 reaction were, 1 x *taq* buffer (Biotech), 1.5 mM $MgCl_2$, 1.25 mM dNTPs, 1 unit *taq*F1 DNA polymerase (Biotech), 10 pmol "primer 1", 10 pmol "primer 2", 5-25 ng DNA, and Milli-Q water to 25 µl. Negative controls of no DNA, no *taq* polymerase, and no primers, along with a positive control, were included in each PCR experiment.

After mixing, the PCR reaction tubes were placed in a thermal cycler (Corbett FTS-960) preheated to 95°C. The samples were then subjected to an initial denaturation step of 3 min at 95°C, then 30 cycles of denaturation at 95°C for 30 sec, annealing at 55-60°C for 30 sec (depending on primers used and specificity wanted), extension at 75°C for 1-7 min (depending on the expected product size), and a final extension at 75°C for 10 min. The reactions were stored at -20°C.

2.14 DNA SEQUENCING.

2.14.1 Manual Sequencing.

Manual sequencing was used for all *tub1* sequencing and, like automatic sequencing, is based on the dideoxy-mediated chain termination method of Sanger *et al.* (1977).

2.14.1.1 Sequencing reactions.

All manual sequencing was performed using the Amplicycle Sequencing Kit (Perkin Elmer) according to the guidelines provided. The label used was [α-^{33}P]dATP.

For each DNA sequencing sample, 2 µl of each of the G,A,T,C termination mixes were aliquoted into four labelled tubes (0.2 ml PCR tubes). A master cocktail containing 20 pmol of primer, ~ 700 ng template DNA, 10 µCi [α-^{33}P]dATP, 4 µl of 10 x cycling mix and Milli-Q water up to 30 µl was prepared and 6 µl of this was added to each of the four tubes (G,A,T,C). Tubes were mixed and immediately placed in a Corbett FTS-960 thermal cycler at 95°C. The samples were then subjected to an initial denaturation step of 2 min at 95°C, then 25 cycles of; 95°C for 30 sec, 60°C for 30 sec, 72°C for 1 min, then 4°C. Upon completion, 4 µl of "stop solution" was added to each tube, and the samples were stored for up to 1 week at-20°C.

Table 2.2 Sequencing and PCR Primers.

Primer.	Size (nt)	Tm* (°C)	Sequence (5` to 3`)	Source/ Reference
pUC/M13 Forward	22	70	GCC AGG GTT TTC CCA GTC ACG A	Perkin Elmer
pUC/M13 Reverse	24	70	GAG CGG ATA ACA ATT TCA CAC AGG	Perkin Elmer
SP6	22	58	TTT AGG TGA CAC TAT AGA ATA C	Promega
T7	23	66	TAA TAC GAC TCA CTA TAG GGC GA	Promega
ITS4	20	58	TCC TCC GCT TAT TGA TAT GC	White *et. al.* (1990)
ITS5	22	62	GGA AGT AAA AGT CGT AAC AAG G	White *et. al.* (1990)
Tub1	23	74	GGG CCA AGG GTC ACT ACA CTG AG	Bidlake (1996)
Tub2	24	72	CGA AGG TAG ACG ACA TCT TGA GAC	Bidlake (1996)
Tub7	22	70	GGG ACA TAC TTG TTG CCG GAC G	This study.
Tub8	21	66	TTG ATG CGT GCC GTA CGC TAG	This study.
Tub9	22	68	ACT TCA CGA CTC ACG ACA CAG C	This study.
Tub10	22	68	TCA TCT CTC GGC GCC ATT GCT A	This study.
151Rep	21	62	GTA GCC TTA CCA TCA ACT GTG	This study.
151Rep2	20	62	GAG CGT TGG ATG TGG ATT GG	This study.
151Rep3	21	64	GAC ATT TTG CGC TGC CTC TTC	This study.
151Rep4	21	62	CTG CAT ACT CGA ACA TCT GAG	This study.
MF4151p1	22	66	ACT TTC AGA TGT CCA TGG CAG C	This study.
MF4151p2	21	62	GTC GCA GTA ATG TCT GAA GAC	This study.
MF4151p3	21	64	GGA CCA GAG GAA CAT ACT TGG	This study.
MF4151p4	21	64	AAC TGC GGT GCT TTA CTG GAG	This study.
MF4151p5	21	64	CAC CAG ACC AAG ATC AAG CAG	This study.
MF4151p6	21	64	AAC ATG ACC TGT CGG GTC ATC	This study.
MF4151p7	21	62	GAC TTG ACA TCT TCC AGC ATC	This study.
152Fep1	21	64	GTC ACG CTG TAG AAG GAC TTG	This study.
152Fep2	21	64	CTG CAT TGC AGC GAT GAT CTG	This study.
152Fep3	20	62	TCT TGT CGA CGA CGG ACT TG	This study.
152Fep4	21	62	CTG AGG ACT ATT TGA GCA CAC	This study.
152Fep5	21	62	GGA GAT TAT TGG GTC TCT AGC	This study.
153Rep	21	68	CTT CGC TGA GCA GAG GGT GAC	This study.
150Fep	22	70	GGA CAT TGG TCG TGT GGT CTG C	This study.
DS150ep1	22	66	CAC AAT CGC ATC AAA CCG TCT C	This study.
DS150ep2	21	62	TTC ATG ATG CGG TCT CCT ATC	This study.
DS150ep3	21	64	GCA GCA AGT CTC TAG TCA TGC	This study.
DS150ep4	20	62	CTA GGA GAG GTT TGC AGC TG	This study.
DS150ep5	21	62	ACA AGA CGG CTC TTA CTG TTC	This study.
DS150ep6	21	64	GAT ACT GCA CGT CGT ATG GTG	This study.

DS150ep7	21	62	TTG ATC GGA GTC AGT ATC GTG	This study.
DST7ep2	21	64	ATC CAG AGC CTA CCT AAG CTC	This study.
DST7ep3	21	62	GTT CGC TTC AAT TCG ATC CAG	This study.
DST7ep4	21	62	CAA CAC GTG CTT CAT CCT ATG	This study.
DST7ep5	21	64	TCT CCT AGC TAC AGC CTA CAC	This study.
DST7ep6	21	62	GGT ACA ACG ATC ATT GTC GAG	This study.
DST7ep7	21	64	ATT CGG CTA CAT GCC CTA CAC	This study.
150Rep	22	70	GCA GAC CAC ACG ACC AAT GTC C	This study.
153Fep	23	70	GTC GAC GGA CAT TAT GGG AGA TG	This study.
152Rep1	20	60	GGT CGT CTG CAA ATA CTA GC	This study.
152Rep2	22	62	GTG TAG CTT TTG CCA TGT ATT G	This study.
152Rep3	22	64	GCT CTG CAT TGC ATC TTA TGT G	This study.
152Rep4	21	66	TGT ACC AGA GCT CGG ACG AAC	This study.
152Rep5	21	64	AGC ATG GCC AAC CAT GGA TAC	This study.
152Rep6	21	64	GTA CGA CAT TGA CGT GCT TGG	This study.
DS152ep	22	66	TGG TCT GGA TGT AGC CAA CAT C	This study.
MF4152p1	21	64	TAC CTG TTC CTG TCT CGT GTC	This study.
MF4152p2	22	64	CTA TCA TTG TCG CTT CGT AAC G	This study.
MF4152p3	21	64	ACA CTG ATT GCG GTA GCG ATG	This study.
MF4152p4	21	66	AGA CCA GCA GGC AGA TGA CAG	This study.
MF4152p5	20	62	ATG GCA CGA GCA GTG ATG AG	This study.
MF4152p6	21	62	GTC TTC AGA CAT TAC TGC GAC	This study.
151Fep	21	62	GTC ATG AGG CAT CCA AAC ATG	This study.
151Fep2	21	60	CGA ACG TCA AAG ACA TTC AAC	This study.
151Fep3	20	62	CGT CTA TGG CCT GAA CTG TC	This study.
151Fep4	20	62	GTA TGC GAG AGC TTC GAA G C	This study.
RTFexon1	21	62	AAC AAA GTC GTC GAG ACC ATC	This study.
RTCexon3	20	62	GAT GCA AGG AAG CAG ACC AC	This study.
RTFintron1	21	60	CTG GTG ATG TAC GTT GTA AAC	This study.

* Calculated as Tm(°C)=2(A+T)+4(G+C) from Itakura *et al.* (1984).

2.14.1.2 PAGE Gel Electrophoresis of Sequencing Reactions.

Sequencing reactions were separated by polyacrylamide gel electrophoresis (PAGE). Sequencing gels were poured with 70 ml of either acrylamide mix (Section) or Long Ranger™ (FMC) mix, polymerised by the addition of 42 µl TEMED (35 µl when using Long Ranger™) and 420 µl of 10% (w/v) ammonium persulphate (350 µl when using Long Ranger™). Once set, the gel was pre-run for 30 min with constant power (65 W) in 1x sequencing TBE buffer (Section 2.4.4). Sequencing reactions were denatured (95°C for 3 min) and 3 µl loaded onto the gel. Once completed, the apparatus was disassembled and the gel dried onto blotting paper at 80°C for 1 hour, then autoradiographed overnight.

2.14.2 Automatic Sequencing.

Automatic sequencing reactions were performed using the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer) with 500 ng template DNA, 3.2 pmol primer, and 8 µl of terminator ready reaction mix in a total volume of 20 µl. Samples were run on an ABI 373 and 377 automatic sequencer.

2.15 ISOLATION OF RNA FROM *D. PINI* CULTURES.

2.15.1 Procedures Used for Treatment of Materials and Reagents for RNA Work.

Glassware was soaked in a chromic acid bath overnight, rinsed, autoclaved and baked for a minimum of 4 hours at 180°C. Previously unopened bags of disposable tips and eppendorfs were autoclaved in RNA treated glassware. RNA solutions were prepared only from stocks dedicated for RNA use. Benches and surfaces were always kept clean by wiping with ethanol and RNase Away (Molecular Bio-Products). Fresh disposable gloves were always worn when handling RNA equipment and during RNA experiments. All solutions were treated with DEPC, except solutions containing Tris-HCl, EDTA, or MOPS buffer (these were made up with DEPC treated water). DEPC was added at 0.01%, incubated overnight at 37°C, and autoclaved twice (or once for 40 min).

2.15.2 Total RNA Isolation.

100-500 mg of *D. pini* mycelium (which was snap frozen in liquid nitrogen upon harvesting) was ground to a fine powder with liquid nitrogen in an ice cold mortar and

pestle. Total RNA was then extracted using TRIzol® Reagent (Life Technologies) as described by the manufacturers. RNA was resuspended in sterile, DEPC treated Milli-Q water. RNA was then quantified by spectrophotometric analysis in a similar manner to DNA quantification, except for RNA an absorbance of 1.0 at 260 nm is equivalent to 40 µg/ml. Purity was indicated by the 260 nm/280 nm ratio, where pure RNA has a ratio of 2.0.

Aliquots of RNA samples were then DNaseI treated using Amplification Grade DNaseI (Life Technologies) according to manufacturers instructions titled "The Use of DNaseI prior to RT-PCR". Reactions were set up to give a final RNA concentration of 100 ng/µl.

2.16 NORTHERN BLOTTING.

2.16.1 Formaldehyde Gel Electrophoresis of RNA.

17.5 ml of formaldehyde and 30 ml of 10 x MOPS buffer was preheated to 55°C. 3 g of agarose was dissolved in 250 ml of nuclease free water and cooled to 55°C. All reagents were then mixed and poured into a gel box immediately.

While the gel was setting, RNA samples were prepared ready for loading. Each sample contained 8 µg total RNA, 2.2 M formaldehyde, 50% (v/v) formamide, and 0.5 x MOPS buffer. Samples were incubated at 55°C for 15 minutes to denature, then 10 x gel loading buffer was added (to give a final concentration of 1 x). Samples were loaded onto the formaldehyde gel and run in 1 x MOPS buffer. An RNA ladder (0.24-9.5 kb; Life Technologies) was run alongside the samples and was prepared in the same way (6 µg of RNA ladder was used). The gel was stained in 5 µg/ml ethidium bromide (in sterile DEPC treated water) and photographed with a ruler as in Section 2.12.1. This method was based on that of Ausubel *et al.*, (1994).

2.16.2 Northern Blotting and Hybridisation

RNA from the formaldehyde gel (Section 2.16.1) was transferred to Hybond-N membrane (Amersham) by capillary blotting as in Section 2.12.1 except the gel was not treated with any solutions, apart from a 20 min wash in water and a 5 min wash in 2 x

SSC. Preparation of the [α-^{32}P]dCTP-labelled DNA probe, hybridisation and autoradiography was performed as in Sections 2.12.2 and 2.12.3

2.17 REVERSE TRANSCRIPTASE PCR (RT-PCR) ANALYSIS OF RNA.

RT-PCR reactions were performed using the SuperScript™ One-Step™ RT-PCR System (Life Technologies) according to the manufacturers instructions. Conditions were optimised with cDNA synthesis being performed at 50°C, each reaction contained 100 ng of DNaseI treated total RNA (Section 2.15.2) and 10 pmol of each primer. Each reaction consisted of 40 PCR amplification cycles, with an annealing temperature of 60°C. Reactions were checked by gel electrophoresis (Section 2.9).

If a PCR product was to be sequenced (Section 2.14.2), it was purified as in Sections 2.6.3 and 2.6.4.

2.18 QUANTIFICATION OF DOTHISTROMIN USING COMPETITIVE ENZYME-LINKED IMMUNOSORBENT ASSAY (ELISA).

Dothistromin levels in broth samples were deduced by competitive ELISA based on the method of Jones *et al* (1993). Standards and samples were prepared by adding 100 µl of each (after particular dilutions) to 100 µl labelled monoclonal antibody (MAb) 10C12A5. All dilutions were performed in a working buffer of: 2 x phosphate-buffered saline (PBS; 8% NaCl, 2.9% Na_2HPO_4, 0.2% KH_2PO_4), 2% chicken egg albumin and 0.2% tween 20. The samples were incubated with MAb at 37°C for one hour (pre-incubation), then 100 µl aliquots from each well were transferred to Doth-MSA microtitre wells and incubated at 37°C for three hours. The plates were then washed six times in PBST (PBS plus 0.1% tween 20), removing any free labelled MAb.

200 µl of freshly prepared substrate mixture (14 mg *o*-phenylene diamine, 0.51 g citric acid, 40 µl of 30% H_2O_2, 27.7 ml of 0.2 M Na_2HPO_4, made to a total volume of 100 ml with water) was added to each micro-titre well, protected from light and incubated with shaking at room temperature for 30 minutes. The peroxidase reaction was stopped by the addition of 50 µl 4 M sulphuric acid to each well. Absorbance was measured at 492 nm in a Dynatech MR 5000 plate reader and the percentage inhibition for each sample and

standard was calculated. A standard curve of percentage inhibition against dothistromin concentration was constructed and unknown concentrations determined.

2.19 PHYLOGENETIC ANALYSIS

Sequences were obtained from the Entrez nucleotide database and aligned with the *D. pini* sequences using Clustal W (Thompson *et. al.*, 1994). Alignments were checked and reformatted in MEGA to nexis files (.nxs) and phylogenetic analysis was performed using the program PAUP* or SplitsTree (Huson *et al.*, 1998).

3. SEQUENCING OF λCGV1 AND IDENTIFICATION OF PUTATIVE DOTHISTROMIN BIOSYNTHETIC GENES.

3.1 INTRODUCTION.

The lambda clone, λCGV1, was previously isolated and the *Sal*I 0.76 kb region shown to contain high sequence identity to the *ver1* and *stcU* genes involved in ST/AF biosynthesis in *A. parasiticus* and *A. nidulans*, respectively (Gillman 1996; Section 1.8). Genes involved in ST/AF biosynthesis are clustered (Sections 1.6.2 and 1.7).

To determine the complete sequence of the *D. pini ver*-like gene, and to determine what other genes surround this region, the complete sequence of λCGV1 was determined along both DNA strands. Sequence analysis identified five open reading frames (ORFs), all of which display significant sequence identity to previously described genes. Three genes (*dkr1*, *dox1* and *dte1*) show similarity to ST/AF biosynthetic cluster genes, indicating that these genes may be invloved in dothistromin biosynthesis. For this reason and to aid discussion, these genes are collectively referred to as dothistromin genes. Southern blot analysis confirmed genomic clustering of the dothistromin genes and copy number.

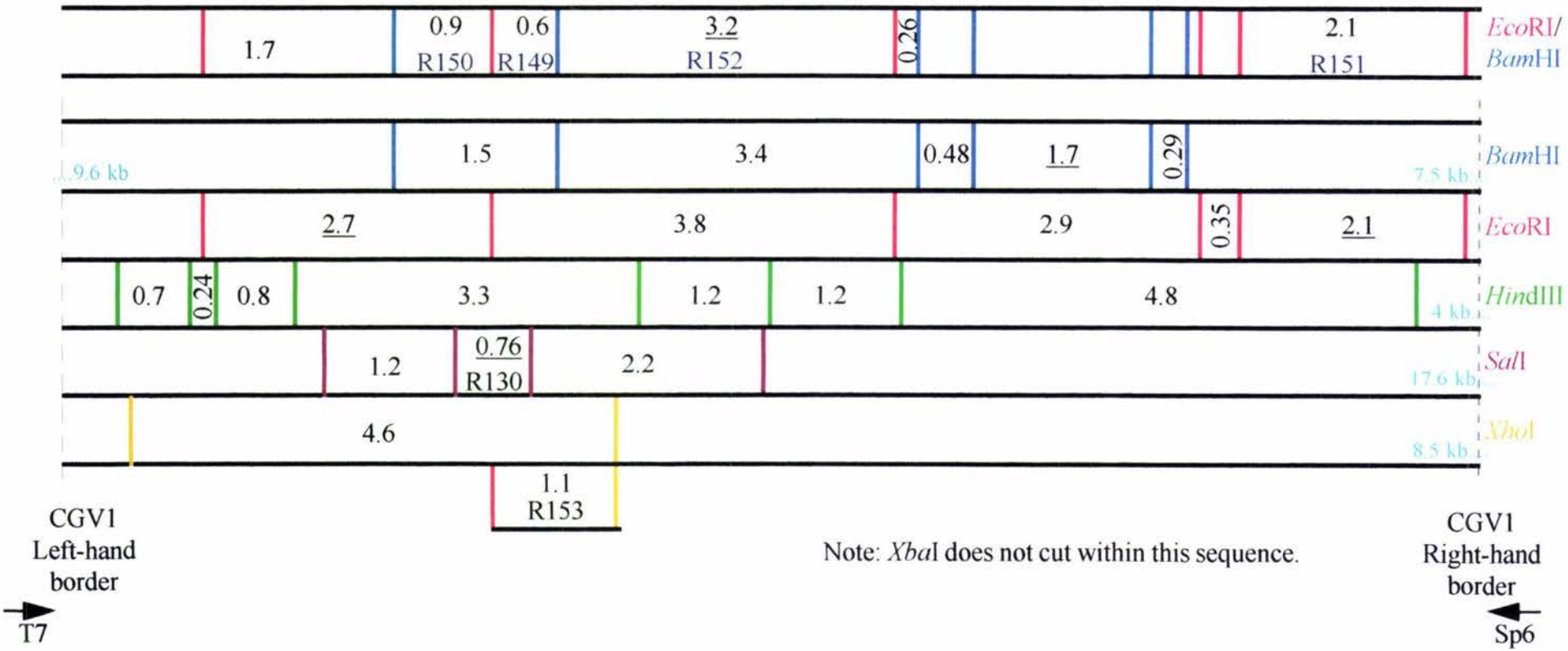
3.2 SUB-CLONING OF λCGV1 FRAGMENTS.

To target regions surrounding the *ver*-like 0.76 kb *Sal*I fragment and to enable efficient sequencing, λCGV1 fragments were sub-cloned into the vector pUC18 (Section 2.11). A preliminary λCGV1 restriction map indicated that four *Eco*RI/*Bam*HI fragments surrounded the *ver*-like sequence (see Figure 3.1 for complete restriction map); these fragments were targetted for "shotgun" cloning.

λCGV1 DNA was doubly digested with *Eco*RI and *Bam*HI (Section 2.8) and then purified for ligation reactions by phenol/chloroform extraction and ethanol precipitation (Section 2.6). Ligation reactions were performed (Section 2.11.1) and transformed into *E. coli* cells by electroporation (Section 2.11.2.1). DNA was isolated from 50 transformants and checked by *Eco*RI/*Bam*HI digestion and gel electrophoresis (Sections 2.5 2.8 and 2.9).

Figure 3.1 λCGV1 Restriction Map.

The λCGV1 restriction map based on the complete nucleotide sequence is shown. Sub-clones (R149 etc.) are indicated. Underlined fragment sizes (all sizes in kb) represent fragments used as probes in hybridisation of Southern blots. Sizes in small font and coloured cyan represent the genomic size of fragments not completely contained on λCGV1 (as determined by Southern analysis, Section 3.10 and Figure 3.14b-f).


1.7
0.9
R150
0.6
R149
3.2
R152
0.26
2.1
R151
EcoRI/
BamHI
1.5
3.4
0.48
1.7
0.29
9.6 kb
7.5 kb
BamHI
2.7
3.8
2.9
0.35
2.1
EcoRI
0.7
0.24
0.8
3.3
1.2
1.2
4.8
4 kb
HindIII
1.2
0.76
R130
2.2
17.6 kb
SalI
4.6
8.5 kb
XhoI
1.1
R153
CGV1
Left-hand
border
T7
Note: XbaI does not cut within this sequence.
CGV1
Right-hand
border
Sp6

This process resulted in the isolation of four clones (Table 2.1): R149 contained a 0.6 kb *Eco*RI/*Bam*HI fragment, R150 contained a 0.9 kb *Eco*RI/*Bam*HI fragment and R152 contained a 3.2 kb *Eco*RI/*Bam*HI fragment. The fourth clone, R151, when digested with *Eco*RI/*Bam*HI produced three bands; linearised vector (pUC18), a 2.1 kb and a 0.26 kb fragment. Digestion with *Eco*RI only, indicated that the 2.1 kb band was an *Eco*RI fragment, implying that the 0.26 kb band was an *Eco*RI/*Bam*HI fragment. It was thought that this clone was an artifact caused by the small fragment being biased in the ligation process, and not the result of a partial digest, as a 0.26 kb band was present in many of the transformants tested (data not shown). This was confirmed by sequencing the R151 insert, which identified a *Eco*RI 2.1 kb fragment and a *Eco*RI/*Bam*HI 0.26 kb fragment; further sequencing of λCGV1 placed the 0.26 *Eco*RI/*Bam*HI fragment elsewhere (Figure 3.1). In future discussion, R151 is referred to as having only the *Eco*RI 2.1 kb fragment present.

The 1.1 kb *Eco*RI/*Xho*I fragment was also sub-cloned as this fragment contains overlapping sequence important for the construction of contigs (Figure 3.1). The resulting plasmid was named R153.

3.3 SEQUENCING AND CONSTRUCTION OF A SINGLE SEQUENCE CONTIG.

3.3.1 Construction of Contig 1 and Contig 2.

Inserts of sub-clones R150, R149, R152, R151 were completely sequenced along both strands and sub-clones R130 and R153 were sequenced to provide overlapping sequence for contig assembly. Initial sequence was obtained using the universal pUC/M13 forward and reverse primers and extended using specifically designed primers (Figure 2.2, Figure 3.2). Sequence was obtained as in Section 2.14 and manipulated using the University of Wisconsin Genetics Computer Group (GCG) software programs.

Direct sequencing, using λCGV1 as template DNA, was performed with the primer DS150ep1 to obtain overlapping sequence 5` to R150, as the preliminary λCGV1 restriction map indicated that R151 was adjacent to R150. However, alignment of these sequences produced two contigs. Contig 1 was 4.7 kb and consisted of sequence from

R150, R130, R149, R153 and R152. Contig 2 containrd sequence from R151 only (2.1 kb). Whilst sequencing these clones, PCR was used to amplify regions not covered by the sub-clones so the complete sequence of λCGV1 could be obtained.

3.3.2 PCR Amplification of Overlapping Regions and Completion of Sequence.

Identification of which restriction fragments were situated between contig 1 and contig 2 could not be resolved by the preliminary restriction map. PCR was used to determine both the orientation and separation of contig 1 and contig 2, and to obtain this "missing fragment". It was thought that this method would be quicker and simpler than Southern blotting, construction of a modified restriction map and cloning of fragments. PCR was performed with λCGV1 as the template DNA and using the same primers used for sequencing (Table 2.2).

Primers closest to the 5` and 3` ends of contig 1 (DS150ep1 and 152Rep2, respectively) were used in PCR reactions with different combinations of primers closest to each end of contig 2 (151Fep2 and 151Rep2). A Mg^{2+} titration was also performed which established the optimal concentration to be 1.5 μM. PCR reactions were performed as described (Section 2.13) with an annealing temperature of 55°C. Products were separated by gel electrophoresis (Section 2.9) and bands were observed in all four lanes (that is, with all four primer combinations), with all but one lane having multiple bands. The lane containing a single product (of 7 kb) was amplified using primers 152Rep2 and 151Rep2 (data not shown). Because bands were observed in each lane the experiment was repeated with the annealing temperature raised to 60°C. Once again a 7 kb product was observed with primers 152Rep2 and 151Rep2. Unfortunately, bands were again observed in the other lanes, however the product sizes had changed from the previous amplification and the bands were faint (data not shown).

To test whether the 7 kb product was specific and not the result of non-specific amplification as observed in the other lanes, nested PCR was performed. The 7 kb product was used as the template DNA, with primers DS152ep and 151Rep4. The resulting product was of the expected size (3.5 kb) and was further evidence that this fragment contained the missing segment between contig 1 and contig 2 (results not shown).

At this stage both contigs had been sequenced along both strands, thus a number of primer options were available and a further experiment was performed to confirm that this fragment was indeed genuine before the effort and expense of cloning and sequencing. Four combinations of primers flanking the "missing fragment" (DS152ep and 151Rep2; 152Rep6 and 151Rep3; 152Rep5 and 151Rep4; DS152ep and 151Rep4) were used in PCR reactions using λCGV1 or *D. pini* genomic DNA as template DNA. In all cases the expected size bands were observed and were the same with either λCGV1 or genomic DNA, indicating the clone and genome organisation are the same (data not shown). To confirm that the same region was amplified in each case, products were purified (Section 2.6.4) and digested with *Eco*RI. All samples exhibited the expected 2.9 kb and 0.35 kb fragments (along with flanking fragments).

This evidence confirmed that, even though bands were observed with other primer combinations, the product originally amplified using primers 152Rep2 and 151Rep2 overlapped contig 1 and contig 2. To clone this region, primers DS152ep and 151Rep3 were used to amplify a 4 kb band. The product was purified, ligated into a pGEM T-vector (Promega) and transformed into *E. coli* by electroporation (Section 2.11.2.1). This primer combination was chosen as it offered a directional cloning strategy if the above strategy did not work. The plasmid containing this 4 kb fragment was called pMF4. This fragment was sequenced along both strands and alignment with contigs 1 and 2 produced a single 10 kb contig (denoted contig 3).

A similar protocol was used to obtain regions between contig 3 and the left and right lambda arms (Figure 3.1). The T7 (primer site in the left λ arm) and SP6 (right λ arm) primers were used in combinations with the contig 3 primers DS150ep and 151Fep5. A 0.8 kb product was produced with primers SP6 and 151Fep5, and a 2.4 kb product with primers DS150ep and T7. These products were also cloned using the T-vector system. The plasmid containing the 0.8 kb product, pMF800, was sequenced and found to contain sequence overlapping R151 and the λ polylinker site, indicating one end of the clone (Figure 3.1 and Figure 3.8). The clone containing the 2.4 kb product was also sequenced, but found to contain only lambda sequence, indicating a non-specific PCR product. The sequence between the left λ arm and contig 3 was therefore obtained by direct

Figure 3.2 Outline of λCGV1 Sequencing.

Schematic outline (BigPicture, GCG) of the λCGV1 sequencing protocol, showing the positions and coverage of primers. The scale is in bp. Primer names are shown on the left. Sequence obtained using the pUC/M13 universial primers is indicated by a sole F (forward primer) or R (reverse) after the sub-clone name/number (151R, pMF800F etc; DS-direct sequence).


151R
pMF800R
151Fep4
pMF800F
151Rep1
151Fep3
151Rep2
151Fep2
151Rep3
151Fep1
151F
151Rep4
MF4152p6
MF4151p1
MF4152p5
MF4151p2
MF4152p4
MF4151p3
MF4152p3
MF4151p4
MF4152p2
MF4151p5
MF4152p1
MF4151p6
DS152ep
MF4151p7
152Rep6
152F
152Rep5
152Fep1
152Rep4
152Fep2
152Fep3
152Rep3
152Rep2
152Fep4
152Fep5
152Rep1
153R
152R
153Rep1
153Fep1
149R
153F
149F
130R
130F
150Rep1
150F
150Fep1
150R
DS150ep1
DST7ep7
DST7ep6
DS150ep2
DST7ep5
DS150ep3
DST7ep4
DS150ep4
DST7ep3
DS150ep5
DST7ep2
DS150ep6
DST7
DS150ep7
CONSENSUS
0
1400
2800
4200
5600
7000
8400
9800
11200
12600

sequencing of λCGV1 using T7 as the starting primer out from the λ arm and DS150ep2 from contig 3. All sequences were aligned and a single contig of 13346 bp was produced, containing 65 sequence fragments. This sequence encompasses the left and right λ arm polylinkers (19 bp each), and 13308 bp of *D. pini* genomic DNA. The sequence strategy is summarised in Figure 3.2. Of the 13.3 kb of sequence, 3.5 kb was obtained from PCR products.

To determine whether PCR amplification of fragments (pMF4, pMF800) had introduced errors, direct sequencing was performed across areas considered important (putative coding sequences). A total of 1.5 kb was re-sequenced and one error was identified.

3.3.3 λCGV1 Contains Part of a Putative Dothistromin Biosynthetic Cluster.

Sequence analysis identified five genes within the 13.3 kb genomic DNA region sequenced (Figure 3.3), all of which display significant identity to previously described genes. Three of these genes show similarity to genes located within the ST and AF biosynthetic gene clusters, indicating that these genes may be involved in dothistromin biosynthesis and that the five genes identified may be part of a dothistromin biosynthetic gene cluster. Inspection of sequence databases was performed using the BLAST search programme (National Centre for Biotechnology Information). Open reading frames (ORFs) were detected using the ORF Finder (NCBI). Sequence analysis of each gene is subsequently discussed in this chapter and is summarised in Table 3.1.

3.4 SEQUENCE OF THE VER-LIKE GENE, *DKR1*.

The *dkr1* (dothistromin ketoreductase) gene displays significant sequence identity to many reductase and dehydrogenase genes, especially ketoreductases involved in ST/AF and melanin biosynthesis. The deduced amino acid sequence of *dkr1* displays; 79% identity (87% similarity) to that of the *A. parasiticus ver1* gene, 78% (88%) with the *A. nidulans stcU* gene, 62% (76%) with the *Cochliobolus heterostrophus brn1* gene, 58% (72%) with the *Magnaporthe grisea thnR* gene and 59% (73%) with the *Colletotrichum lagenarium thr1* gene (functions and references for these genes are shown in Table 3.1).

The *dkr1* ORF is 916 bp in length (792 bp without intron sequences, Section 3.4.1), encoding a deduced polypeptide product of 263 amino acid residues, which is

Figure 3.3 The Putative Dothistromin Gene Cluster.

A schematic map of the five genes identified in the λCGV1 clone is shown. Arrows indicate postion size, and direction of each gene. Nucleotide positions are in bp. *Eco*RI (E), *Bam*HI (B) and *Xho*I (X) restriction sites are indicated. The GC content of specific regions (promoter region- top figure, coding region- middle, terminator- bottom) is indicated and positions of probes used in Southern analysis are shown.

GC Content.

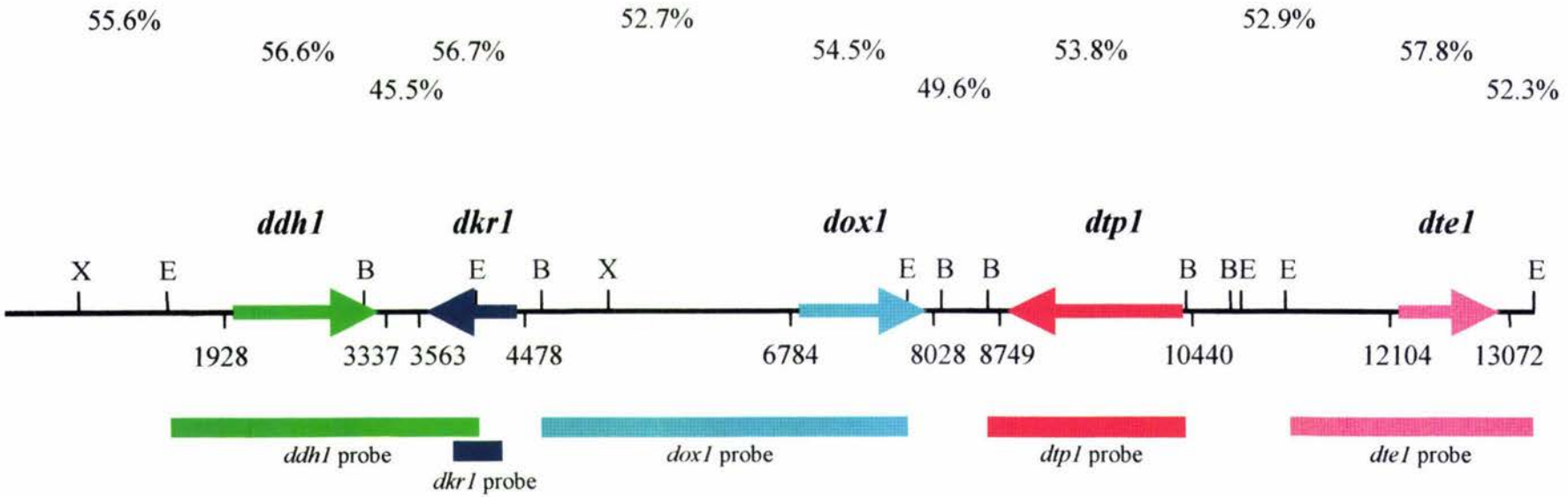

55.6%
56.6%
45.5%
56.7%
52.7%
54.5%
49.6%
53.8%
52.9%
57.8%
52.3%
ddh1
dkr1
dox1
dtp1
dte1
X
E
B
E
B
X
E
B
B
B
BE
E
E
1928
3337
3563
4478
6784
8028
8749
10440
12104
13072
ddh1 probe
dkr1 probe
dox1 probe
dtp1 probe
dte1 probe

Table 3.1 Summary of *D. pini* Genes present in the λCGV1 Clone.

Genes and proteins displaying the highest identity and similarity to the deduced amino acid sequences of the five genes in the putative dothistromin gene cluster are shown. MAD- UDP-N-acetyl-D-mannosaminuronic acid dehydrogenase. * indicates that value is the percentage identity and similarity to thioesterase domain only.

D. pini gene.	Similar gene/protein	Putative function	% identity (% similarity)	Reference (or accession number)
dkr1	*ver1- Aspergillus parasiticus*	ketoreductase- AF biosynthesis	79 (87)	Skory *et al.*, 1992
	stcU- A. nidulans	ketoreductase- ST biosynthesis	78 (88)	Keller *et al.*, 1994; Brown et al.,
	brn1- Cochliobolus heterostrophus	reductase- melanin biosynthesis	62 (76)	1996
	thr1- Colletotrichum lagenarium	reductase- melanin biosynthesis	59 (73)	Shimizu *et al.*, 1997
	thnR- Magnaporthe grisea	reductase- melanin biosynthesis	58 (72)	Perpetua *et al.*, 1996
				Vidal-Cros *et al.*, 1994
dox1	*stcC- A nidulans*	Oxidase- ST biosynthesis	16 (29)	Brown *et al.*, 1996
	chloroperoxidase gene- *Caldariomyces*	Chloroperoxidase- biosynthesis of chlorinated secondary	21 (31)	Nuell *et al.*, 1988
	fumago	metabolites		
dte1		Thioesterase domains of:		
	stcA- A. nidulans	PKS- ST biosynthesis	22 (38)*	Yu and Leonard, 1995
	pksA- A. parasiticus	PKS- AF biosynthesis	28 (45)*	Chang *et al.*, 1995
	pksP- A. fumigatus	PKS- conidial pigmentation biosynthesis	39 (58)*	Y17317
	pks1- C. lagenarium	PKS- Melanin biosynthesis	27 (42)*	Takano *et al.*, 1995
dtp1	*toxA- C. carbonum*	HC-toxin efflux pump	24 (41)	Pitkin *et al.*, 1996
	C3H1 protein- *Schizosaccharomyces pombe*	Intergral membrane transporter	25 (45)	Z68144
	SPBC protein- *S. pombe*	Intergral membrane transporter	25 (41)	AL021748
ddh1	*wbpA- Pseudomomas aeruginosa*	dehydrogenase- B-band lipopolysaccharide biosynthesis	21 (38)	Burrows et al., 1996
	cap5O- Staphylococcus aureus	dehydrogenase- common-antigen biosynthesis	18 (32)	Kiser and Lee, 1998
	MAD- *Methanococcus jannaschii*	deydrogenase.	18 (32)	U67494

comparable with the gene products described above. The translation initiation codon (ATG) was determined at nucleotide position 4478 and was chosen as it was followed by the longest ORF and fitted in context with the predicted amino acid sequences of the other ketoreductase genes. The *dkr1* translation initiation sequence (CTCCCATAATGTC) is consistent with the consensus sequence (Kozak sequence) for filamentous fungi [CNNNCA(A/C)NATGGC] (Bruchez *et al.*, 1993). The putative termination codon (TAG) is at nucleotide position 3565. The nucleotide sequence and deduced amino acid sequence for *dkr1* is shown in Figure 3.4.

3.4.1 The *dkr1* Gene Contains Two Introns.

Putative intervening regions were identified by the presence of 5` [GT(A/G/T)NGTY] and 3` (YAG) fungal consensus intron splice site sequences (Ballance, 1986; Bruchez *et al.*, 1993), as well as maintenance of amino acid similarity with the other sequences. Two putative introns were identified, the first 66 nt in length and starting at λCGV1 nucleotide position 4148, and the second 58 nt long and starting at position 3893 (Figure 3.4). Intron positions were confirmed by the direct sequencing of a RT-PCR product which was amplified from 100 ng total *D. pini* RNA using primers RTFexon1 and RTCexon3 (data not shown). The position of the first intron is conserved between *dkr1* and the ST/AF genes *stcU* and *ver1*, but differs from the melanin genes *brn1* and *thr1* (the positions of introns in these two genes are identical). The position of the second intron in *dkr1* is unique. No comparisons were made with *thnR* as only the cDNA sequence was available.

3.4.2 Amino Acid Sequence Comparison.

The deduced amino acid sequence for *dkr1* was aligned with those of *ver1*, *stcU*, *thnR*, *thr1* and *brn1* using the GCG PILEUP program (Figure 3.5). This identified high identity between the genes throughout the sequence, except for a uniform lack of identity near the middle of the alignment (between DKR1 E^{71} and Q^{81}; Figure 3.5).

The N-terminal of DKR1 contains a putative NADPH binding site (^{19}GXGIGX^{24}) which is conserved in all six genes. Analysis indicated that the *M. grisea* THNR reductase protein is a class-B dehydrogenase (Vidal-Cros *et. al.*, 1994). Due to the extent of sequence identity observed between the proteins (as mentioned above), it is highly likely

Figure 3.4 Nucleotide and Deduced Amino Acid Sequence of the *Dothistroma pini dkr1* Gene.

The coding sequence, promoter and terminator regions and deduced amino acid sequence of *dkr1* are shown. Numbers on the right refer to the positions of nucleotides, relative to the start of the the λCGV1 sequence (top), and amino acid residues (bottom, red). Initiation and termination codons are shown in green and intron consensus sequences are shown in blue with intron sequences in lower case. Locations of primers used in RT-PCR are indicated. The putative NADPH binding site is shown in magenta and underlined, and the putative active site is coloured cyan. The 3` end of *ddh1* is also indicated.

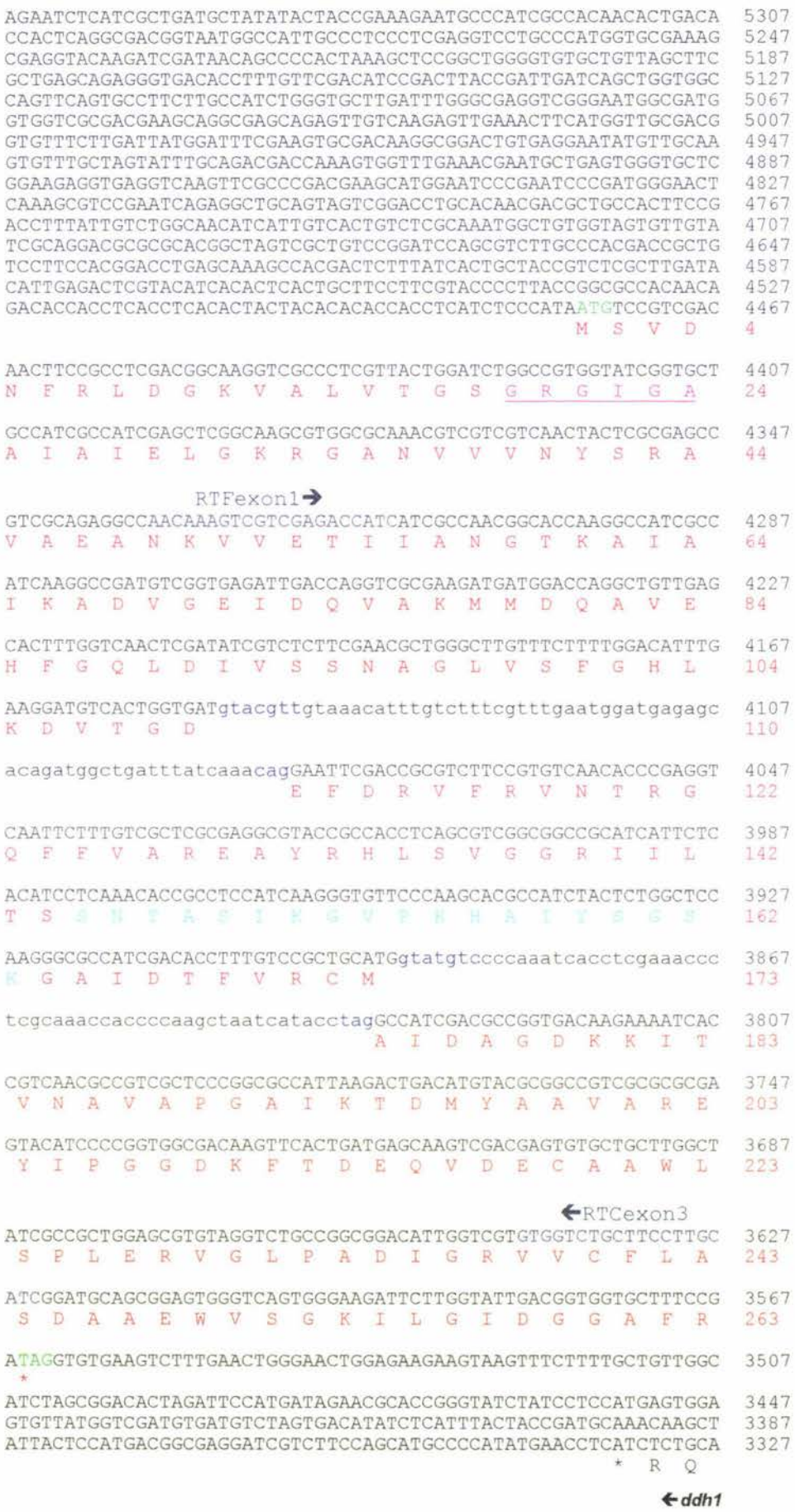
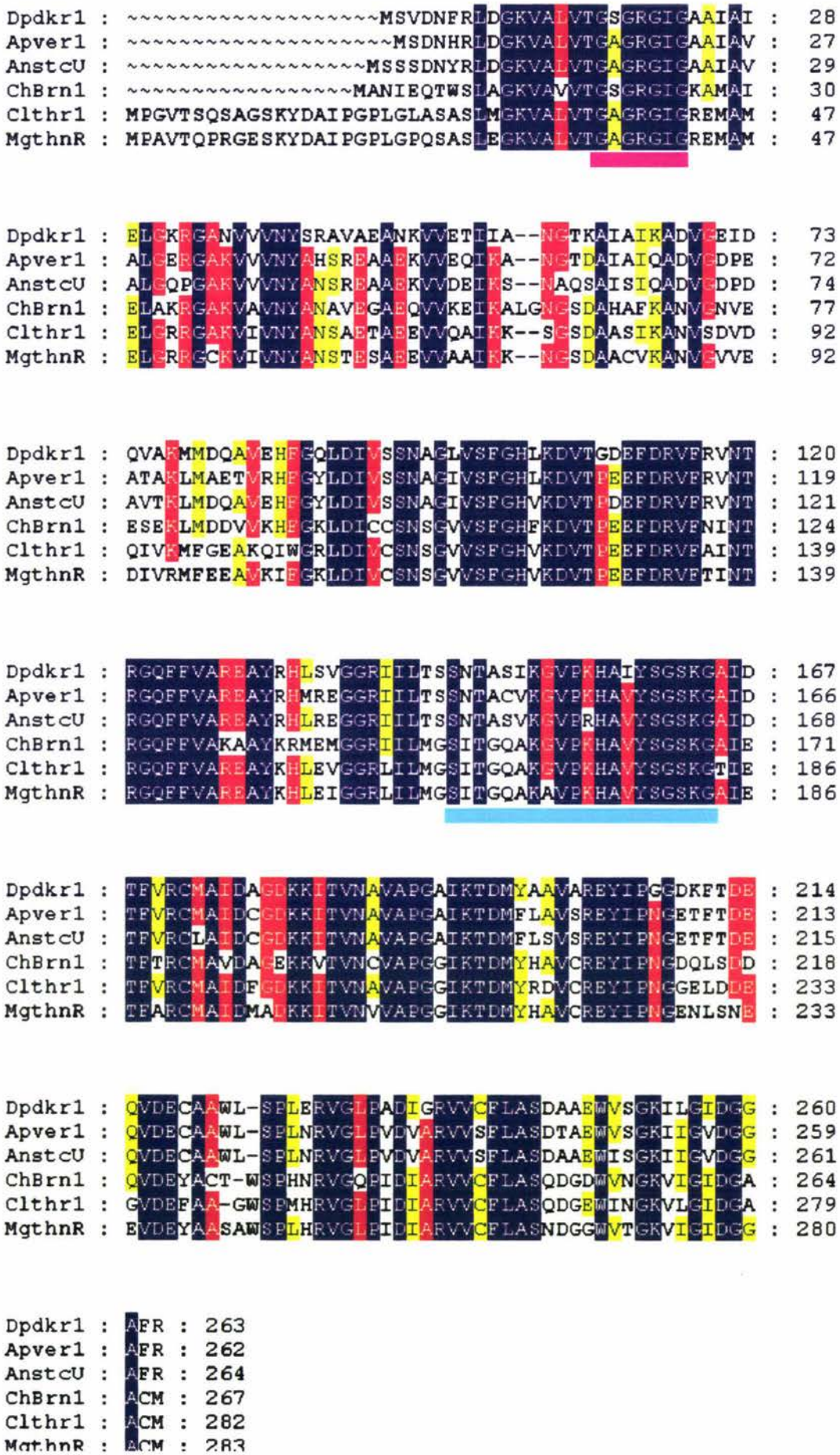

RTFexon1
RTCexon3
ddh1

Figure 3.5 Alignment of the Deduced Amino Acid Sequence of *dkr1* With Other Reductases.

The deduced amino acid sequence of the *D. pini dkr1* gene (Dpdkr1, Figure 3.4) was aligned with that of the *A. parasiticus ver1* gene (Apver1), *A. nidulans stcU* (AnstcU), *C. heterostrophus brn1* (ChBrn1), *C. lagenarium thr1* (Clthr1), and the *M. grisea thnR* gene (MgthnR). See Table 3.1 for references. Sequence alignment was performed with the GCG PILEUP program using the default settings. Amino acid sequence identity is indicated by shading: dark blue indicates the residue is conserved in all six sequences, red indicates the residue is present in five of the six sequences and residues shaded yellow are conserved in four of the six sequences. The putative NADPH binding site is indicated by a magenta-coloured line and the putative active site is indicated by a cyan-coloured line.


Dpdkr1 : ~~~~~~~~~~~~~~~~~~~~MSVDNFRLDGKVALVTGSGRGIGAAIAI : 28
Apver1 : ~~~~~~~~~~~~~~~~~~~~MSDNHRLDGKVALVTGAGRGIGAAIAV : 27
AnstcU : ~~~~~~~~~~~~~~~~~~~~MSSSDNYRLDGKVALVTGAGRGIGAAIAV : 29
ChBrn1 : ~~~~~~~~~~~~~~~~~~~~MANIEQTWSLAGKVAVVTGSGRGIGKAMAI : 30
Clthr1 : MPGVTSQSAGSKYDAIPGPLGLASASLMGKVALVTGAGRGIGREMAM : 47
MgthnR : MPAVTQPRGESKYDAIPGPLGPQSASLEGKVALVTGAGRGIGREMAM : 47
Dpdkr1 : ELGKRGANVVVNYSRAVAEANKVVETIIA--NGTKAIAIKADVGEID : 73
Apver1 : ALGERGAKVVVNYAHSREAAEKVVEQIKA--NGTDAIAIQADVGDPE : 72
AnstcU : ALGQPGAKVVVNYANSREAAEKVVDEIKS--NAQSAISIQADVGDPD : 74
ChBrn1 : ELAKRGAKVAVNYANAVEGAEQVVKEIKALGNGSDAHAFKANVGNVE : 77
Clthr1 : ELGRRGAKVIVNYANSAETAEEVVQAIKK--SGSDAASIKANVSDVD : 92
MgthnR : ELGRRGCKVIVNYANSTESAEEVVAAIKK--NGSDAACVKANVGVVE : 92
Dpdkr1 : QVAKMMDQAVEHFGQLDIVSSNAGLVSFGHLKDVTGDEFDRVFRVNT : 120
Apver1 : ATAKLMAETVRHFGYLDIVSSNAGIVSFGHLKDVTPEEFDRVFRVNT : 119
AnstcU : AVTKLMDQAVEHFGYLDIVSSNAGIVSFGHVKDVTPDEFDRVFRVNT : 121
ChBrn1 : ESEKLMDDVVKHFGKLDICCSNSGVVSFGHFKDVTPEEFDRVFNINT : 124
Clthr1 : QIVEMFGEAKQIWGRLDIVCSNSGVVSFGHVKDVTPEEFDRVFAINT : 139
MgthnR : DIVRMFEEAVKIFGKLDIVCSNSGVVSFGHVKDVTPEEFDRVFTINT : 139
Dpdkr1 : RGQFFVAREAYRHLSVGGRIILTSSNTASIKGVPKHAIYSGSKGAID : 167
Apver1 : RGQFFVAREAYRHMREGGRIILTSSNTACVKGVPKHAVYSGSKGAID : 166
AnstcU : RGQFFVAREAYRHLREGGRIILTSSNTASVKGVPRHAVYSGSKGAID : 168
ChBrn1 : RGQFFVAKAAYKRMEMGGRIILMGSITGQAKGVPKHAVYSGSKGAIE : 171
Clthr1 : RGQFFVAREAYKHLEVGGRLILMGSITGQAKGVPKHAVYSGSKGTIE : 186
MgthnR : RGQFFVAREAYKHLEIGGRLILMGSITGQAKAVPKHAVYSGSKGAIE : 186
Dpdkr1 : TFVRCMAIDAGDKKITVNAVAPGAIKTDMYAAVAREYIPGGDKFTDE : 214
Apver1 : TFVRCMAIDCGDKKITVNAVAPGAIKTDMFLAVSREYIPNGETFTDE : 213
AnstcU : TFVRCLAIDCGDKKITVNAVAPGAIKTDMFLSVSREYIPNGETFTDE : 215
ChBrn1 : TFTRCMAVDAGEKKVTVNCVAPGGIKTDMYHAVCREYIPNGDQLSDD : 218
Clthr1 : TFVRCMAIDFGDKKITVNAVAPGGIKTDMYRDVCREYIPNGGELDDE : 233
MgthnR : TFARCMAIDMADKKITVNVVAPGGIKTDMYHAVCREYIPNGENLSNE : 233
Dpdkr1 : QVDECAAWL-SPLERVGLPADIGRVVCFLASDAAEWVSGKILGIDGG : 260
Apver1 : QVDECAAWL-SPLNRVGLPVDVARVVSFLASDTAEWVSGKIIGVDGG : 259
AnstcU : QVDECAAWL-SPLNRVGLPVDVARVVSFLASDAAEWISGKIIGVDGG : 261
ChBrn1 : QVDEYACT-WSPHNRVGQPIDIARVVCFLASQDGDWVNGKVIGIDGA : 264
Clthr1 : GVDEFAA-GWSPMHRVGLPIDIARVVCFLASQDGEWINGKVLGIDGA : 279
MgthnR : EVDEYAASAWSPLHRVGLPIDIARVVCFLASNDGGWVTGKVIGIDGG : 280
Dpdkr1 : AFR : 263
Apver1 : AFR : 262
AnstcU : AFR : 264
ChBrn1 : ACM : 267
Clthr1 : ACM : 282
MgthnR : ACM : 283

that the the *D. pini* DKR1 reductase (and indeed all aligned sequences) belongs to the same class. DKR1 also contains a putative active site ($^{145}SX_{13}YSGSK^{163}$); this motif is conserved in all the ketoreductase sequences aligned. This motif is consistent with the $SX_{10\text{-}30}YXXXK$ sequence which has been shown to be the active site of the *Drosophila* alcohol dehydrogenase (Chen *et. al.*, 1993).

3.5 *DOX1* ENCODES A PUTATIVE HEME-BINDING OXIDASE.

The *dox1* (dothistromin oxidase) gene displays similarities to oxidase, peroxidase and chloroperoxidase genes of fungi and bacteria. The deduced amino acid sequence of *dox1* displays 21% identity (31% similarity) to that of the *Caldariomyces fumago* (*Leptoxyphium fumago*) chloroperoxidase gene and 16% (29%) to the *A. nidulans* oxidase gene, *stcC*, which is present in the ST cluster (functions and references for these genes are shown in Table 3.1).

The *dox1* ORF is 1245 bp in length, encoding a predicted polypeptide product of 414 amino acid residues, which is 90-100 residues larger than the products of the chloroperoxidase and *stcC* genes. The translation initiation codon was determined at nucleotide position 6784 as this is the start of the *dox1* ORF; the next in-frame start codon, 234 nucleotides downstream, was discounted as it is after a region with considerable homology including the putative heme-binding site. The termination codon (TAG) is at nucleotide position 8026. The nucleotide sequence and deduced amino acid sequence for *dox1* is shown in Figure 3.6.

The *dox1* coding sequence contains no obvious intron consensus sequences which is consistent with the *A. nidulans stcC* gene and the *C. fumago* chloroperoxidase gene which do not contain introns.

The deduced amino acid sequence of *dox1* was aligned with that of the *A. nidulans stcC* gene and the *C. fumago* chloroperoxidase gene (Figure 3.7). The *C. fumago* chloroperoxidase is a heme-containing glycoprotein and has a well characterised heme-binding active site, ^{49}PCPALNALANHG^{60} (Blanke and Hager, 1988; Blanke and Hager, 1990). This active site is conserved in DOX1 and the region contains the highest level of identity present throughout the alignment (Figure 3.7), suggesting that *dox1* is a heme-binding oxidase. The *C. fumago* chloroperoxidase protein contains a 21 amino acid signal

Figure 3.6 Nucleotide and Deduced Amino Acid Sequence of the *Dothistroma pini dox1* Gene.

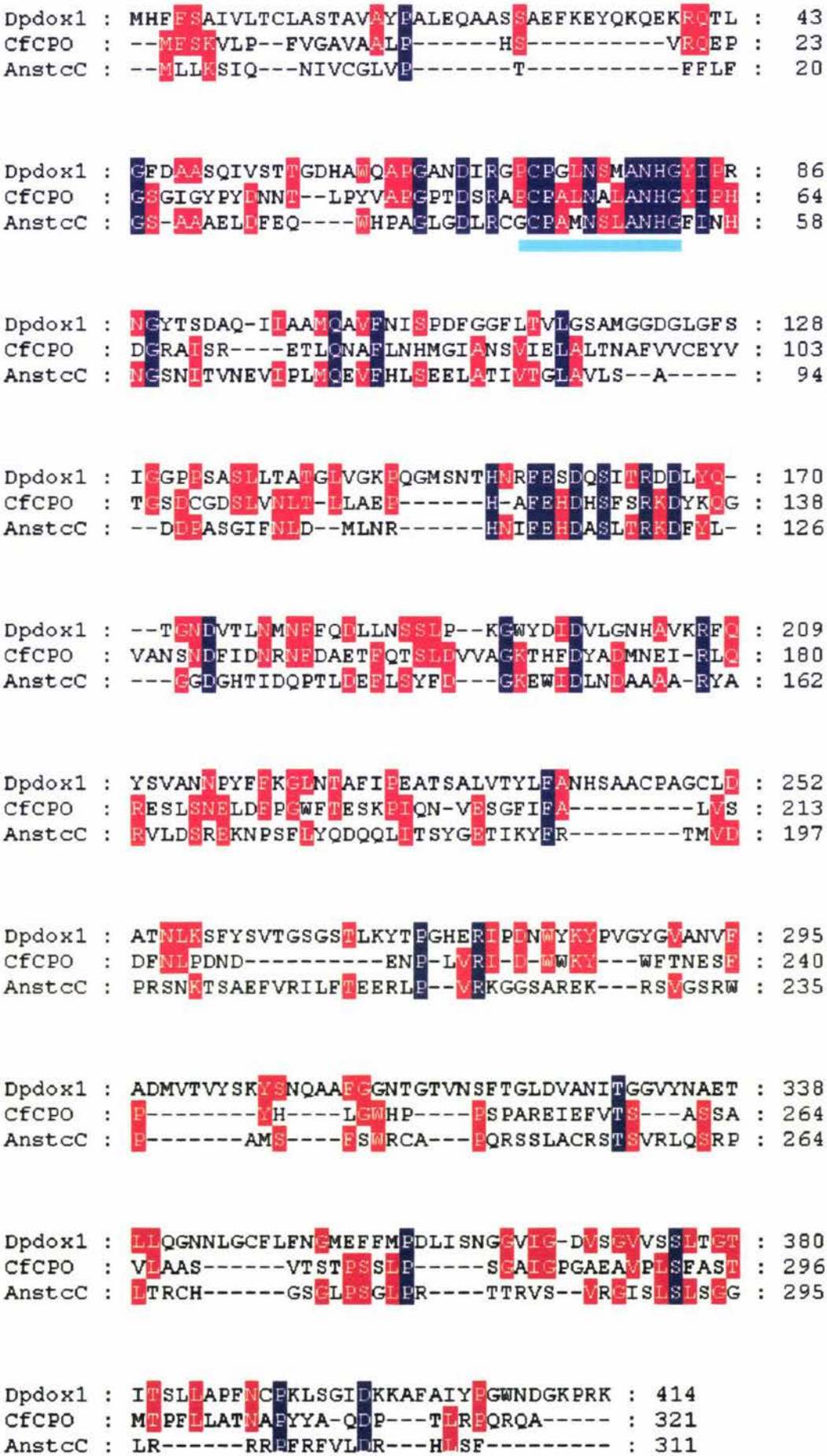
The coding sequence, promoter and terminator regions and deduced amino acid sequence of *dox1* are shown. Numbers on the right refer to the positions of nucleotides, relative to the start of the λCGV1 sequence (top), and amino acid residues (bottom, red). The putative heme-binding site is coloured cyan. Initiation and termination codons are shown in green. Locations of primers used in RT-PCR are indicated. TATA promoter motifs are shown in blue.

```
CAGCCTTTTGTCCCACCCGTGAAGATCATCCCGCTTATCCTCGTATTCACGGGCTCTGCA  5940
TTGCATCTTATGTGATCCCGGCTCGCGAACAGCGCCAGACACGATGCCCAGATCCATGAT  6000
GCTTTGTTTACGATTGCATCTAATCTAGCAGGACTCTGTAGTGTGCTCAAATAGTCCTCA  6060
GCCCAGACCAGTGTTTCGCGGTCTTGCACCAACTGTCCGTCCACTTTGGCTGAGCCACAT  6120
CTGCGTTTCTTTGGTACGATTCGGCAATCCTGGAATCTAGTATTGCCATCACCGACCGGC  6180
GGATGGAGGTGGTTGGCTCCGCGATAGGTTATGCCCATCGCTTACTTGGTCGATGGCTTG  6240
CACCAGTGCCTTCAAGAGGTACTATCTCGAAGTCCGCTCGCTTGCGATGTGCTCATATTG  6300
TTGGCAGGATTACAGGGCTGTGAGTTCACGAGCCCAGTGTCTTAGATCGAACCAATTGTT  6360
CTTTTAACCGGGTCTGGATAATGCCCGATGCCACAGTCTTGTCTTGATGGCTCCGGTGTG  6420
ACTCAGCCGTAAAGGCGGAGTTGCTTGAGGACGCAGCATGTACCAGAGCTCGGACGAACA  6480
ACGCTACTCGTCCTTGAAAATGAATGGCCACAGCGTGCTGGTAAGACGTCCGATGGCTTG  6540
GCCGCTTTGACATCGAATGTCAGCCACGTTTTCGGAAATTCTGTACAGGTGCCTCGTTAT  6600
TGGTTGCCAAGTCCGTCGTCGACAAGATCAGGCGACGAGTATGGCTTGCGCCGTATTACA  6660
AGGCGACGGACGACGGAGCTCAGCAGCAAGGACGGTATTATAAGAGAGCTTCATCGTACC  6720
GCAAAGCTTCACAGAATCCGCCAAGCAAAGGATCAACCTGTACAAGCCGACATCTAACCC  6780
AGCATGCATTTCTTCTCCGCAATCGTTCTGACCTGCTTGGCATCTACGGCGGTGGCATAC  6840
   M  H  F  F  S  A  I  V  L  T  C  L  A  S  T  A  V  A  Y    19

CCTGCGTTGGAACAAGCCGCATCATCTGCAGAGTTCAAGGAATACCAGAAGCAAGAGAAG  6900
P  A  L  E  Q  A  A  S  S  A  E  F  K  E  Y  Q  K  Q  E  K    39

CGTCAGACTCTCGGCTTCGATGCTGCTTCTCAAATCGTCAGCACCACTGGCGACCATGCT  6960
R  Q  T  L  G  F  D  A  A  S  Q  I  V  S  T  T  G  D  H  A    59

TGGCAAGCACCAGGAGCAAATGACATTCGTGGGCCATGCCCAGGCTTGAACAGCATGGCC  7020
W  Q  A  P  G  A  N  D  I  R  G  P  C  P  G  L  N  S  M  A    79

 152Rep5→
AACCATGGATACATTCCGCGCAACGGATACACCTCCGATGCGCAGATCATCGCTGCAATG  7080
N  H  G  Y  I  P  R  N  G  Y  T  S  D  A  Q  I  I  A  A  M    99

CAGGCAGTTTTCAACATCTCCCCAGACTTTGGCGGCTTCTTGACCGTCCTTGGTTCGGCC  7140
Q  A  V  F  N  I  S  P  D  F  G  G  F  L  T  V  L  G  S  A   119

ATGGGTGGCGATGGTTTGGGCTTCTCCATCGGCGGACCACCTTCAGCGTCGTTGTTGACG  7200
M  G  G  D  G  L  G  F  S  I  G  G  P  P  S  A  S  L  L  T   139

GCCACGGGTCTCGTCGGCAAACCTCAGGGCATGAGCAACACCCACAACCGCTTCGAAAGC  7260
A  T  G  L  V  G  K  P  Q  G  M  S  N  T  H  N  R  F  E  S   159

GACCAGAGCATCACGCGCGATGATCTGTACCAGACTGGCAACGATGTGACCTTGAACATG  7320
D  Q  S  I  T  R  D  D  L  Y  Q  T  G  N  D  V  T  L  N  M   179

AACTTCTTCCAGGACCTCCTCAATTCGTCATTGCCAAAAGGATGGTACGACATTGACGTG  7380
N  F  F  Q  D  L  L  N  S  S  L  P  K  G  W  Y  D  I  D  V   199

CTTGGAAATCATGCAGTGAAGAGGTTCCAATACTCCGTGGCGAACAATCCTTACTTCTTC  7440
L  G  N  H  A  V  K  R  F  Q  Y  S  V  A  N  N  P  Y  F  F   219

AAAGGTCTCAACACTGCCTTCATTCCGGAAGCGACATCGGCACTTGTTACGTACCTCTTC  7500
K  G  L  N  T  A  F  I  P  E  A  T  S  A  L  V  T  Y  L  F   239

                                                  ←152Fep1
GCCAACCACTCGGCCGCTTGCCCAGCGGGCTGCTTGGATGCCACCAATCTCAAGTCCTTC  7560
A  N  H  S  A  A  C  P  A  G  C  L  D  A  T  N  L  K  S  F   259

TACAGCGTGACCGGCTCCGGCTCAACCCTTAAATACACCCCGGGCCACGAACGCATCCCC  7620
Y  S  V  T  G  S  G  S  T  L  K  Y  T  P  G  H  E  R  I  P   279

GACAATTGGTACAAATACCCCGTAGGATATGGCGTCGCCAACGTCTTCGCCGACATGGTG  7680
D  N  W  Y  K  Y  P  V  G  Y  G  V  A  N  V  F  A  D  M  V   299

ACAGTATACTCCAAGTACTCCAACCAAGCAGCATTCGGTGGGAACACGGGTACCGTCAAC  7740
T  V  Y  S  K  Y  S  N  Q  A  A  F  G  G  N  T  G  T  V  N   319

AGCTTTACTGGTCTGGATGTAGCCAACATCACCGGCGGTGTCTACAACGCCGAGACCCTT  7800
S  F  T  G  L  D  V  A  N  I  T  G  G  V  Y  N  A  E  T  L   339

CTGCAAGGCAACAACTTGGGATGCTTCCTGTTCAACGGGATGGAATTCTTCATGCCGGAT  7860
L  Q  G  N  N  L  G  C  F  L  F  N  G  M  E  F  F  M  P  D   359
```

```
CTCATCTCGAACGGCGGAGTCATCGGCGATGTATCGGGTGTGGTGTCGAGCTTGACGGGG  7920
L  I  S  N  G  G  V  I  G  D  V  S  G  V  V  S  S  L  T  G    379

ACTATCACCTCGCTACTGGCACCGTTCAACTGTCCGAAGCTTTCTGGGATTGATAAGAAG  7980
T  I  T  S  L  L  A  P  F  N  C  P  K  L  S  G  I  D  K  K    399

GCGTTCGCCATCTATCCTGGCTGGAATGATGGCAAGCCAAGGAAGTAGGGGTCCAAGTCA  8040
A  F  A  I  Y  P  G  W  N  D  G  K  P  R  K  *                414

GTATGACAGGCGTTGGTTGTAGGACCTCGTGATTAGAGAAACCTGGCGGGTTGTGGCGAA  8100
TCAGGATCCAATGGGCGTTGAGAGCCTTGCGACGTCGTACATCTTGTGCTGTTTCACCAG  8160
TCTTTACGGTCAGCATGAAGATACACATGAATCCATTTCTTGGAGTCTGCATCGTGCGGT  8220
TGTACCTGTTCCTGTCTCGTGTCTTTGGCCGCCGTTGTGACATGTATTGATAGACATGAA  8280
AGATGCTGGAAGATGTCAAGTCCATCAATGGAAGCATTTCTGGGCTGCGCGGAGTCGATG  8340
GAGAGTCGATGGCTCCGAAGGTGGTCGATGAGCAGAAGATTGTTTGGTGGAGACGCGAGG  8400
```

Figure 3.7 Alignment of the Deduced Amino Acid Sequence of *dox1* and Other Oxidases.

The deduced amino acid sequence of the *D. pini dox1* gene (Dpdox1, Figure 3.6) was aligned with that of the *C. fumago* chloroperoxidase gene (CfCPO) and the *A. nidulans stcC* gene (AnstcC). See Table 3.1 for references. Sequence alignment was performed with the GCG PILEUP program using the default settings. Amino acid sequence identity is indicated by shading: dark blue indicates the residue is conserved in all three sequences and red indicates the residue is present in two of the three sequences. The putative heme binding site is indicated by a cyan-coloured line.


Dpdox1 : MHFFSAIVLTCLASTAVAYPALEQAASSAEFKEYQKQEKRQTL : 43
CfCPO : --MFSKVLP--FVGAVAALP------HS----------VRQEP : 23
AnstcC : --MLLKSIQ---NIVCGLVP-------T----------FFLF : 20
Dpdox1 : GFDAASQIVSTTGDHAWQAPGANDIRGPCPGLNSMANHGYIPR : 86
CfCPO : GSGIGYPYDNNT--LPYVAPGPTDSRAPCPALNALANHGYIPH : 64
AnstcC : GS-AAAELDFEQ----WHPAGLGDLRCGCPAMNSLANHGFINH : 58
Dpdox1 : NGYTSDAQ-IIAAMQAVFNISPDFGGFLTVLGSAMGGDGLGFS : 128
CfCPO : DGRAISR----ETLQNAFLNHMGIANSVIELALTNAFVVCEYV : 103
AnstcC : NGSNITVNEVIPLMQEVFHLSEELATIVTGLAVLS--A----- : 94
Dpdox1 : IGGPFSASLLTATGLVGKPQGMSNTHNRFESDQSITRDDLYQ- : 170
CfCPO : TGSDCGDSLVNLT-LLAEP------H-AFEHDHSFSRKDYKQG : 138
AnstcC : --DDPASGIFNLD--MLNR------HNIFEHDASLTRKDFYL- : 126
Dpdox1 : --TGNDVTLNMNFFQDLLNSSLP--KGWYDIDVLGNHAVKRFQ : 209
CfCPO : VANSNDFIDNRNFDAETFQTSLDVVAGKTHFDYADMNEI-RLQ : 180
AnstcC : ---GGDGHTIDQPTLDEFLSYFD---GKEWIDLNDAAAA-RYA : 162
Dpdox1 : YSVANNPYFFKGLNTAFIPEATSALVTYLFANHSAACPAGCLD : 252
CfCPO : RESLSNELDFPGWFTESKPIQN-VESGFIFA---------LVS : 213
AnstcC : RVLDSREKNPSFLYQDQQLITSYGETIKYFR--------TMVD : 197
Dpdox1 : ATNLKSFYSVTGSGSTLKYTPGHERIPDNWYKYPVGYGVANVF : 295
CfCPO : DFNLPDND----------ENP-LVRI-D-WWKY---WFTNESF : 240
AnstcC : PRSNKTSAEFVRILFTEERLP--VRKGGSAREK---RSVGSRW : 235
Dpdox1 : ADMVTVYSKYSNQAAPGGNTGTVNSFTGLDVANITGGVYNAET : 338
CfCPO : P--------YH----LGWHP----PSPAREIEFVTS---ASSA : 264
AnstcC : P-------AMS----FSWRCA---PQRSSLACRSTSVRLQSRP : 264
Dpdox1 : LLQGNNLGCFLFNGMEFFMPDLISNGSVIG-DVSGVVSSLTGT : 380
CfCPO : VLAAS------VTSTPSSLP-----SGAIGPGAEAVPLSFAST : 296
AnstcC : LTRCH------GSGLPSGLPR----TTRVS--VRGISLSLSGG : 295
Dpdox1 : ITSLLAPFNCPKLSGIDKKAFAIYPGWNDGKPRK : 414
CfCPO : MTPFLLATNAPYYA-QDP---TLRPQRQA----- : 321
AnstcC : LR------RRPFRFVLDR---HLSF--------- : 311

peptide which has an atypical glutamate (E) residue at the signal peptide processing site (Nuell *et al.*, 1988). Although this region is not conserved in DOX1, the N-terminal is hydrophobic, a property which is characteristic of signal peptides. Therefore, DOX1 may contain a signal peptide.

3.6 *DTE1* IS A PUTATIVE THIOESTERASE GENE.

The *dte1* (dothistromin thioesterase) gene displays similarities to the thioesterase domain of polyketide synthases (PKS) involved in ST/AF and conidial pigmentation biosynthesis (see Table 3.1).

The translation start site of *dte1* is difficult to predict. There are two in-frame ATG codons within 69 bp of each other and the 21 amino acids between the ATG codons contain no highly conserved regions. The Kozak consensus sequence for these start codons are similar. Depending on the translation start site, the ORF is predicted to encode a product of 322 or 300 amino acids. The sequence and deduced amino acid sequence for *dte1* is shown in Figure 3.8.

The *dte1* gene displays a distinct ORF, implying the *dte1* gene may comprise part of a type II PKS. Alignment of the deduced amino acid sequence of *dte1* with the thioesterase domains of STCA, PKSA, PKSP and PKS1 displayed some regions of identity and a complete conservation of the thioesterase active site motif ^{140}GPYXXXGXSXG^{150} (Figure 3.9; Yu and Leonard, 1995). Thioesterase releases the polyketide product from the enzyme complex.

3.7 *DTP1* IS A PUTATIVE DOTHISTROMIN TOXIN PUMP.

The *dtp1* (dothistromin toxin pump) gene displays similarities to many small efflux pumps present in fungi and bacteria that belong to the drug resistance protein family of the major facilitator superfamily of transmembrane proteins. The deduced amino acid sequence of *dtp1* displays 24% identity (41% similarity) to the *Cochliobolus carbonum* HC-toxin efflux pump encoded by *toxA*, as well as similarities to two uncharacterised *Schizosaccharomyces pombe* integral membrane transporter proteins, C3H1 (25% identity, 45% similarity) and SPBC (25%, 41%) (functions and references for these genes are shown in Table 3.1).

Figure 3.8 Nucleotide and Deduced Amino Acid Sequence of the *Dothistroma pini dte1* Gene.

The coding sequence, promoter and terminator regions and deduced amino acid sequence of *dte1* is shown. Numbers on the right refer to the positions of nucleotides, relative to the start of the λCGV1 sequence (top), and amino acid residues (bottom, red; numbering starts from the first ATG). The two possible translation start sites are indicated (in green) and the amino acid sequence between the two is underlined. Nucleotide sequence in blue and double underline indicates the right λ-arm of λCGV1. The putative thioesterase active site motif is coloured cyan. Locations of primers used in RT-PCR are indicated and a possible TATA motif is shown in blue.

```
ACCCGTCTGGCACGCGTTCAGACGGTACCCATGCTGGCTGTGCCTCCGCCTCCCAACGCA  11280
GAATGGCAGCTGAAGCCGTTCATCCTGACCCATGTGATATCGAGACGTGAGGTCTGACCC  11340
ACGATGGAATATTGACATGCCCCAGAATCGTGACCATGGGAAGTGCACCGAGTCATGAGG  11400
CATCCAAACATGTGCGCCCTCGCAGTACCTGGCCGATACGTGCGATGTGTGCAAGGGCGA  11460
TGTGACTAATGCATCCTAGGCTTCGAACGATGTGTGGATCTTCGAGGACTTCATCACGGG  11520
TGCCTGGTCAATTATTCAAGACAAAATCACCTGCGAACAGACATGCACGATGGCAAACGA  11580
TGTAGTAGGGGCTCCGGCGTGGGATAAGCTCAGATGTTCGAGTATGCAGAGCGTTGTCCC  11640
GTCAATGCCAAGCGAGATCAGCGTGCATCTCACTGACAGCGAGACACCACTGAACGAGCA  11700
GAATCCACGCACACATCGAGTCGCATCGTAAGATCGATTGTCAATGAGCACCACGATGAT  11760
GATGGTCAAGGCATAGACTCGCGAGCAGAGCCAAGTTCATGATGACATGTTGGTCACAGA  11820
GCATTATCGGAGTCCGACAACACCTGGTGGAGATTCTACTAAAGGTCAAGCCAGATTTAC  11880
GCCCGCCGCAACATAGCTTGATCCTTCCGCTACTCAGTGACGACGACGGTTCTGCACGGC  11940
TGCAGGGGGCGTCGCCGAACGTCGAAGACATTCAACACCGACAACGCAGCAGACAGACGC  12000
ATTTTGACAGATATACGAAGCAACGCCAAGCCACGAGAGGAATGCTCCATTCGATGAAAC  12060
CTCGGCGAACATAGCAAATCGTCTAGCGAAGAGGCAGCGCAAAATGTCTGCGGCCGTGAC  12120
                                              M  S  A  A  V  T    6
ATCGGCCGCTTCCATCGGCGTCCCCAAAGCACCACACCCAGCCCTCAACATGGACATCCC  12180
 S  A  A  S  I  G  V  P  K  A  P  H  P  A  L  N  M  D  I  P   26

GGCACTGGACGAGAAACAGGTCAAACGTGCCCTCATCGCCGCCGATCACGATTCCAAGAA  12240
 A  L  D  E  K  Q  V  K  R  A  L  I  A  A  D  H  D  S  K  K   46

GAAATTGAAGCGCTCCAGATCCTCACAATCACCATCAACACATCCACCCGCAACCTCAGT  12300
 K  L  K  R  S  R  S  S  Q  S  P  S  T  H  P  P  A  T  S  V   66

CCTCCTCCAAGGCAATCCCCGCCAAGCCACAAAGAACCTCTTCCTCTTCCCCGACGGCTC  12360
 L  L  Q  G  N  P  R  Q  A  T  K  N  L  F  L  F  P  D  G  S   86

CGGTGCCGCAGCCTCCTACACCCACCTCACCCTCATCTCTCGCGACCTCGCCGTCTATGG  12420
 G  A  A  A  S  Y  T  H  L  T  L  I  S  R  D  L  A  V  Y  G   106

 151Fep3→
CCTGAACTGTCCCTACCTCCGCTCCCCCCAGGACTGGAAATGCGGCCCCCAAGACCTGAC  12480
 L  N  C  P  Y  L  R  S  P  Q  D  W  K  C  G  P  Q  D  L  T   126

CCCTCTCTTCATCTCAGAGATCCAACGCCGGCAACCTTCCGGACCTTACTACATAGGCGG  12540
 P  L  F  I  S  E  I  Q  R  R  Q  P  S  G  P  Y  Y  I  G  G   146

CTACTCCACCGGCGGAATCGCAGCCTTTGACGCTGCCCAAGCTCTCGACAAACTGGGTGA  12600
 Y  S  T  G  G  I  A  A  F  D  A  A  Q  A  L  D  K  L  G  E   166

                                          ←151Rep2
GAAAGTCGAGAGACTTATCCTCATCGACTCCCCCTGTCCAATCCACATCCAACGCCTCCC  12660
 K  V  E  R  L  I  L  I  D  S  P  C  P  I  H  I  Q  R  L  P   186

CAGCAGATTAATGGACTACCTCAAACGCGTCCACGCCTTCAACTCCCGCGGCCGCCCACC  12720
 S  R  L  M  D  Y  L  K  R  V  H  A  F  N  S  R  G  R  P  P   206

CCCGGCATGGGTCTTCGAACACTTCGAAGCCAACACGACCAATCTCCAAAAATACCGCCC  12780
 P  A  W  V  F  E  H  F  E  A  N  T  T  N  L  Q  K  Y  R  P   226

ACGGCCCTTTGAAGCTTACAAAGAACCGCGCACCCACATCATCTATGCTCGTCAGGGCGT  12840
 R  P  F  E  A  Y  K  E  P  R  T  H  I  I  Y  A  R  Q  G  V   246

ATGCGAGAGCTTCGAAGCTGGTGTCCCGCAGATGGAAATCTTGGAAGAGGACCCGAAGGA  12900
 C  E  S  F  E  A  G  V  P  Q  M  E  I  L  E  E  D  P  K  E   266

GATGAAGTGGATTATGTGTGCACGGAGTGATTTTGGTCCGCTCGGGTGGGAGAAGTTGTT  12960
 M  K  W  I  M  C  A  R  S  D  F  G  P  L  G  W  E  K  L  L   286

GAATGAGGAGGAGATTTTTGTCGAGATTGTGGAGGGGGCGAATCATTTTGGCATGATGAG  13020
 N  E  E  E  I  F  V  E  I  V  E  G  A  N  H  F  G  M  M  R   306

GGGGGATGCCGCCGAGAGGTTGGCAGGCTGTATTGGTAGGGCGGTCGCATGAAGTCGATG  13080
 G  D  A  A  E  R  L  A  G  C  I  G  R  A  V  A  *            322

ATGGGTGAGTGATGATTGGCAATGATTCATGATGGGGTCAAGGCGTTGAGAAGCGATTTC  13140
ATGGTCAGTGTGCTTGCTGAGGAGTGGGACGTGTTTCGTGTGTGTGTTTCGGCTTGTAGC  13200
GGGAAATTGCGATGATGAGTCTGCGACGGCTATGGCACAGTTGATGGTAAGGCTACTTTG  13260
GCTCCGGGAATTCATGAACGCGAGTACGAGGGGTGGGTCGGCTAGGATTGCCAAGCACTC  13320
GTCAAGATCGAGAATTCCCGGGATCC                                     13346
```

Figure 3.9 Alignment of the Deduced Amino Acid Sequence of *dte1* with PKS Thioesterase domains.

The deduced amino acid sequence of the *D. pini dte1* gene (Dpdte1, Figure 3.8) was aligned with the thioesterase domains of: *A. nidulans* STCA (AnstcA), *A. parasiticus* PKSA (AppksA), *A. fumigatus* PKSP (AfpksP) and *C. lagenarium* PKS1 (Clpks1). See Table 3.1 for references. Sequence alignment was performed with the GCG PILEUP program using the default settings. Amino acid sequence identity between the sequences is indicated by shading: dark blue indicates the residue is conserved in all five sequences, red indicates the residue is present in four of the five sequences and residues conserved in three of the five sequences are shaded yellow. The putative thioesterase active site is indicated by a cyan-coloured line.

Note that, to increase the alignment accuracy, the first 1820 amino acids have been deleted for AnstcA, Appks1, AfpksP and Clpks1. Therefore, the numbers which represent residue positions are incorrect. To obtain the correct position, the number should be added to 1820.

```
Dpdte1 : ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ :   -
AnstcA : IEPLDLGDELFRNVLRIVSEESGVALDELSAETVFADIGIDSL :  43
Appks1 : ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ :   -
AfpksP : ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ :   -
Clpks1 : ~~~~~~~~~~~~~~~~~~ALGMDSLMSLSILGTLREKSGQDIP :  25

Dpdte1 : ~~~~~~~~~~~~MSAAVTSAASIGVPKAPHPALNMDIPALDEK :  31
AnstcA : SSMVITSRFREDLGMSLDSSFNLFEEVPTVARLQEFFGTTSGS :  86
Appks1 : ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ :   -
AfpksP : ~~~~~~~~~~~~~~~~~~~MDLPAEFFLENPTLDAVQAALDLK :  24
Clpks1 : NDLFVTNPSLLEVEKALGIGPKPKPAAAPKPAKSA--PAASRR :  66

Dpdte1 : -QVKRALIAADHDSKKKLKRSRSSQSPS-----THPPATSVLL :  68
AnstcA : TTGSSGSGSSEDETDSIPSTPEAYT-TADTRVPECRPTTSVVL : 128
Appks1 : ~~~~~~IFASEDHGHSSESGADTGSPPALDLKPYCRPSTSVVL :  37
AfpksP : PKMVPAATPVSEPIRLLETIDNTKPKTS-----RHPPATSILL :  62
Clpks1 : EKVEPTKEINTHPGNTTASITKPPPPTEIIDNYPHRKATSILL : 109

Dpdte1 : QGNPRQATKNLFLFPDGSGAAASYTHLTLISRDLAVYGLNCPY : 111
AnstcA : QGLPQMAKQILFMLPDGGGSASSYLTIPRLHADVAIVGLNCPY : 171
Appks1 : QGLPMVARKTLFMLPDGGGSAFSYASLPRLKSDTAVVGLNCPY :  80
AfpksP : QGNPHTATKKLFMFPDGSGSASSYATIPALSPDVCVYGLNCPY : 105
Clpks1 : QGSTRTATKNLWMVPDGSGCATSYTEISQVSSNWAVWGLFSPF : 152

Dpdte1 : LRSPQDWKCGPQDLTPLFISEIQRRQPSGPYYICGYSTGGIAA : 154
AnstcA : ARDPENMNCTHQSMIQSFCNEIKRRQPEGPYHLGGWSSGGAFA : 214
Appks1 : ARDPENMNCTHGAMIESFCNEIRRRQPRGPYHLGGWSSGGAFA : 123
AfpksP : MKTPQNLTCSLDELTEPYLAEIRRRQPKGPYSFGGWSAGGICA : 148
Clpks1 : MKTPEEYKCGVYGMAAKFIEAMKARQSKGPYSLAGWSAGGVIA : 195

Dpdte1 : FDAA-QALDKLGEKVERLILIDSPCPIHIQRLPSRLMDYLKRV : 196
AnstcA : Y-VTAEALINAGNEVHSLIIIDAPVPQVMEKLPTSFYEYCNNL : 256
Appks1 : Y-VVAEALVNQGEEVHSLIIIDAPIPQAMEQLPRAFYEHCNSI : 165
AfpksP : FDAARQLILEEGEEVERLLLLDSPFPIGLEKLPPRLYKFFNSI : 191
Clpks1 : YEIVNQ-LTKAGETVENLIIIDAPCPVTIEPLPRSLHAWFASI : 237

Dpdte1 : HAFNS---R----GRPPPAWVFEHFEANTTNLQKYR--PRPFE : 230
AnstcA : GLFSNQPGGTTDGTAQPPPYLIPHFQATVDVMLDYRVAPL--- : 296
Appks1 : GLFATQPGASPDGSTEPPSYLIPHFTAVVDVMLDYKLAPL--- : 205
AfpksP : GLFGD---G----KRAPPDWLLPHFLAFIDSLDAYKAVPLPFN : 227
Clpks1 : GLLGE---GDDEAAKKIPSWLLPHFAASVTALSNYTAEPIPKE : 277

Dpdte1 : ----AYKEPRTHIIYARQGVCESFEAGVPQ-MEILEEDPKEMK : 268
AnstcA : ---KTNRMPKVGIIWASETVMDEDNAPK----------MKGMH : 326
Appks1 : ---HARRMPKVGIVWAADTVMDERDAPK----------MKGMH : 235
AfpksP : DSKWAKKMPKTYLIWAKDGVCGKPGDPRPEPAEDGSEDPREMQ : 270
Clpks1 : ------KCPNVMAIWCEDGVCHLPTDPRPDPYPTG-----HAL : 309

Dpdte1 : WIMCARSDFGPLGWEKLLNEEEI-FVEIVEGANHFGMMRGDAA : 310
AnstcA : FMVQKRWDFGPDGWDVVCPGAVFDILR-AEGANHLR~~~~~~~ : 361
Appks1 : FMIQKRTEFGPDGWDTIMPGASFDIVR-ADGANHFTLMQKEHV : 277
AfpksP : WLLNDRTDLGPNKWDTLVGPQNIGGIHVMEDANHFTMTTGQKA : 313
Clpks1 : FLLDNRTDFGPNRWDEYLDVNKFRTRHM--PGNHFSMMHGDYV : 350

Dpdte1 : ERLAGCIGRAVA~~~~~ : 322
AnstcA : ~~~~~~~~~~~~~~~~~ :   -
Appks1 : SIISDLIDRVMA~~~~~ : 289
AfpksP : KELSQFMATAMSS~~~~ : 326
```

Figure 3.10 Nucleotide and Deduced Amino Acid Sequence of the *Dothistroma pini dtp1* Gene.

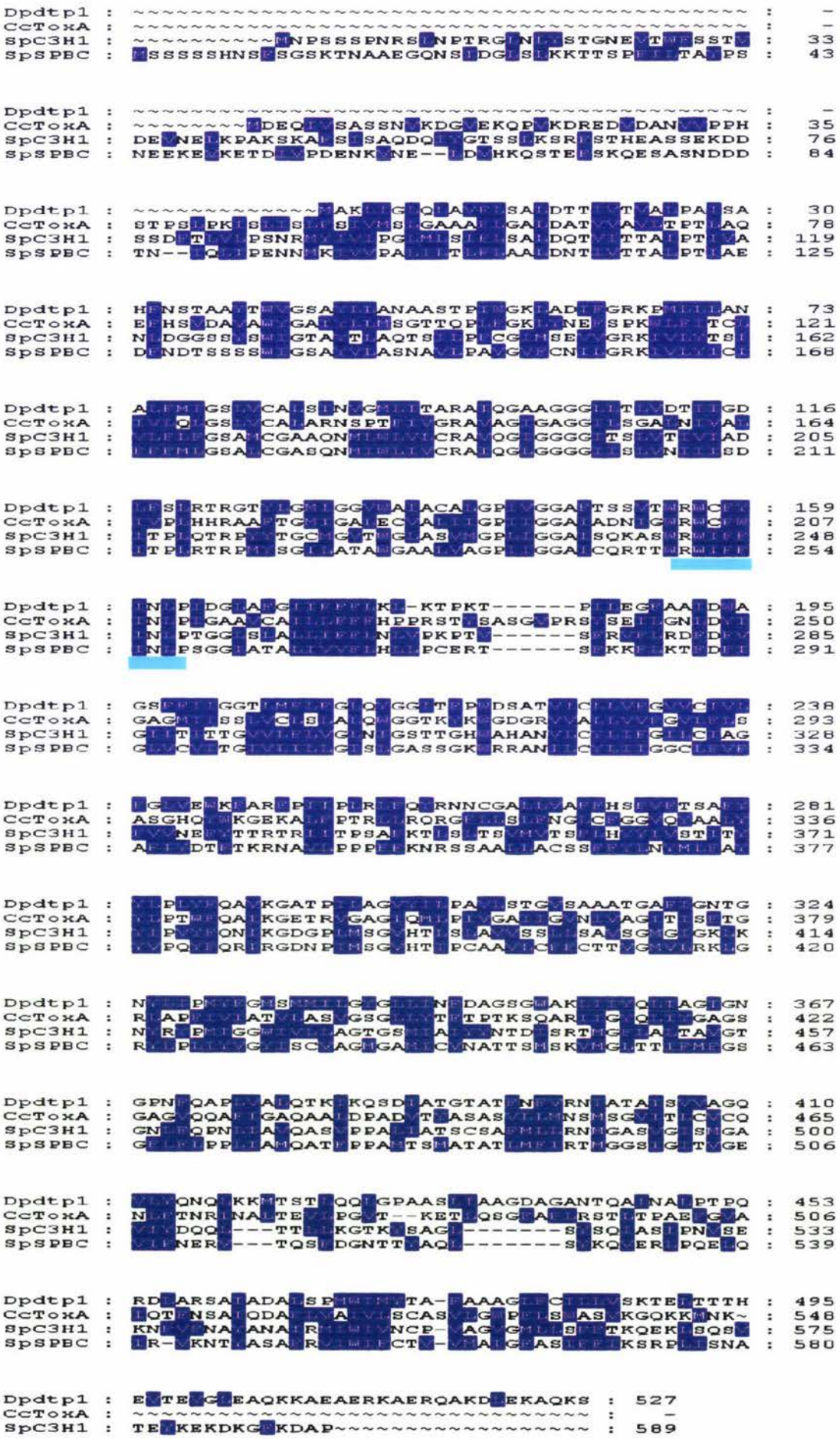
The coding sequence, promoter and terminator regions and deduced amino acid sequence of *dtp1* is shown. Numbers on the right refer to the positions of nucleotides, relative to the start of the λCGV1 sequence (top), and amino acid residues (bottom, red). Initiation and termination codons are shown in green while intron consensus sequences are shown in blue with intron sequences in lower case. Locations of primers used in RT-PCR are indicated. Possible TATA and CAAT promoter motifs are shown in blue. A tentative polyadenlyation signal is shown in bold.

```
TCTGGGGCATGTCAATATTCCATCGTGGGTCAGACCTCACGTCTCGATATCACATGGGTC  11307
AGGATGAACGGCTTCAGCTGCCATTCTGCGTTGGGAGGCGGAGGCACAGCCAGCATGGGT  11247
ACCGTCTGAACGCGTGCCAGACGGGTTCCTCTCAGCTCGTCGATCCTTTCATGTCAACAA  11187
GCGAGTCCGCACACCGGACTTTCAGATGTCCATGGCAGCCTTGGCAAAATGCGATGGAAT  11127
TCCTAAGGATTTCATGTCGTTTTGGATTCTGAAATCCGACGTGTCCGCTCATTCTCATCC  11067
CGGCGATCCGAGCGTCTGAGACTACGCCATAAGGGTATTAGCGTCACCGCTCAGCTGGCT  11007
CTCAACCAACCATCGGCTCTGTGGCGCGACCCGATCCCTACCGCAACCTTCGATGACCCT  10947
AGATCGAGATGGTGAAGGCCGACAGTGTCGGTCTTGACATTCCAACCAATCCCGATTCAG  10887
CACGAGCAGTTCGCAGCAGCTGGACATATCTTTCCCCCCTTGCAAGAATGTCTCTTGGGG  10827
CGAGATATTTCGGTACTGGTTGTTGTGAGATGACGCGCGTGAAATTAGAATTCCGCTATA  10767
AGGGAGCTAGCTGTATCCTTGGGGGGTGTCGCTGAAGTAGTAAGGAATTGCAGAGGCTAC  10707
GACCTGGCAATCTTCTCTTGGCAGAGAAATACAGACAACAAACAAGTCGCAGTAATGTCT  10647
GAAGACCACACCAAAGCCGACAACTTGTCCGAGAAGGATCCGCATTCACCAGAGAGGAGT  10587
GATAGTAGTAGTCATGAAGATGCTCACGCACGGGAGGAGGAGGAGAGCTCAGATGACGAT  10527
GGAGCGCTGGATGGCAAGCCAGCGAGCTTGATTGCGATTGTGATGATTGCATTGTCGGTC  10467
AGTCTTCTTCACGATGAGCATGCGATATGGCAAAGCTGATAGGTCTCCAGCTGGCAGTAT  10407
                          M  A  K  L  I  G  L  Q  L  A  V  F   12

TCCTCTCAGCACTCGACACAACCATCGTCACCGTGGCTCTACCAGCCATCTCAGCACATT  10347
  L  S  A  L  D  T  T  I  V  T  V  A  L  P  A  I  S  A  H  F   32

TCAACTCCACCGCGGCGTACACATGGGTCGGATCCGCATATCTACTTGCCAACGCAGCTT  10287
  N  S  T  A  A  Y  T  W  V  G  S  A  Y  L  L  A  N  A  A  S   52

CAACACCGATCTGGGGAAAGCTCGCCGACATCTTCGGACGCAAGCCGATGCTTCTCCTCG  10227
  T  P  I  W  G  K  L  A  D  I  F  G  R  K  P  M  L  L  L  A   72

CCAATGCGTTGTTCATGATTGGCTCACTTGTCTGTGCGCTTTCGATCAACGTTGGCATGC  10167
  N  A  L  F  M  I  G  S  L  V  C  A  L  S  I  N  V  G  M  L   92

TCATCACTGCTCGTGCCATCCAAGGTGCCGCAGGTGGTGGTCTGTTGACCTTGGTTGATA  10107
  I  T  A  R  A  I  Q  G  A  A  G  G  G  L  L  T  L  V  D  T  112

                                                MF4151P3➔
CTATCATTGGCGATCTGTTCTCCCTCCGGACCAGAGGAACATACTTGGGTATGATCGGTG  10047
  I  I  G  D  L  F  S  L  R  T  R  G  T  Y  L  G  M  I  G  G  132

GAGTTTGGGCCATCGCTTGCGCCCTTGGCCCGATCGTTGGAGGTGCCTTTACTTCTAGTG  9987
  V  W  A  I  A  C  A  L  G  P  I  V  G  G  A  F  T  S  S  V  152

TGACCTGGAGATGgtaagtactatcacgcgctacatcatggcaagaacagcaaatggtaa  9927
  T  W  R  W                                                   156

catactttagGTGCTTCTATATCAACCTGCCGATTGATGGACTCGCTTTTGGCATCATCT  9867
           C  F  Y  I  N  L  P  I  D  G  L  A  F  G  I  I  F   173

TCTTCTTCTTGAAACTGAAGACCCCAAAGACGCCTATACTTGAGGGCTTTGCTGCCATCG  9807
  F  F  L  K  L  K  T  P  K  T  P  I  L  E  G  F  A  A  I  D  193

ACTGGGCCGGTAGCTTCTTCgtatgtattcgccgatacagttcgacagcaagcagactga  9747
  W  A  G  S  F  F                                             199

cattttcgcagATCATCGGCGGCACTCTTATGTTTCTCTTCGGACTCCAATATGGCGGCA  9687
           I  I  G  G  T  L  M  F  L  F  G  L  Q  Y  G  G  I   216

TCACCTTCCCATGGGACTCGGCGACTGTCATCTGCCTGCTGGTCTTCGGCGTAGTCTGCA  9627
  T  F  P  W  D  S  A  T  V  I  C  L  L  V  F  G  V  V  C  I  236

TTGTCCTCTTCGGTTTGGTCGAGTGGAAGTTCGCCAGATTTCCAATCATCCCACTGCGCC  9567
  V  L  F  G  L  V  E  W  K  F  A  R  F  P  I  I  P  L  R  L  256

TGTTCCAGTACCGAAACAACTGCGGTGCTTTACTGGTGGCTTTCTTCCACTCCTTCGTCT  9507
  F  Q  Y  R  N  N  C  G  A  L  L  V  A  F  F  H  S  F  V  F  276

TCACCTCGGCTTTCTACTACTTGCCTTTGTACTTCCAGGCCGTCAAAGGAGCGACGCCAA  9447
  T  S  A  F  Y  Y  L  P  L  Y  F  Q  A  V  K  G  A  T  P  I  296

TTCTGGCTGGTGTCTACATCCTGCCAGCTGTGTTGTCCACAGGTGTCAGCGCTGCAGCAA  9387
  L  A  G  V  Y  I  L  P  A  V  L  S  T  G  V  S  A  A  A  T  316

CGGGAGCGTTCATCGGAAACACCGGCAACTACCTCATCCCCATGTACTTTGGCATGAGTA  9327
  G  A  F  I  G  N  T  G  N  Y  L  I  P  M  Y  F  G  M  S  M  336

TGATGATCCTCGGCTACGGTCTGCTCATCAACTTCGACGCTGGCTCCGGCTGGGCGAAGC  9267
  M  I  L  G  Y  G  L  L  I  N  F  D  A  G  S  G  W  A  K  L  356
```

```
  I  I  Y  Q  L  I  A  G  I  G  N  G  P  N  F  Q  A  P  L  V  376

TTGCACTCCAGACCAAGATCAAGCAGAGCGATATCGCAACCGGCACAGCCACTTTCAACT  9147
  A  L  Q  T  K  I  K  Q  S  D  I  A  T  G  T  A  T  F  N  F  396


                ←MF4152P3
TTGTACGCAACATCGCTACCGCAATCAGTGTCGTGGCTGGCCAAGTCCTTTATCAGAACC  9087
  V  R  N  I  A  T  A  I  S  V  V  A  G  Q  V  L  Y  Q  N  Q  416

AGCTTAAGAAGATGACCTCTACTCTGCAGCAGCTTGGTCCAGCAGCGTCGCTGATTGCCG  9027
  L  K  K  M  T  S  T  L  Q  Q  L  G  P  A  A  S  L  I  A  A  436

CAGGTGATGCCGGCGCCAACACCCAGGCGATCAACGCCCTACCTACACCGCAGAGAGACC  8967
  G  D  A  G  A  N  T  Q  A  I  N  A  L  P  T  P  Q  R  D  L  456

TTGCAAGATCAGCCATTGCGGATGCACTGTCGCCCATGTGGATCATGTACACGGCTTTTG  8907
  A  R  S  A  I  A  D  A  L  S  P  M  W  I  M  Y  T  A  F  A  476

CAGCGGCAGGACTGTTCTGTATCCTGCTCGTCAGCAAGACTGAGTTGACAACGACCCACG  8847
  A  A  G  L  F  C  I  L  L  V  S  K  T  E  L  T  T  T  H  E  496

AAGTGACTGAGGTCGGCCTCGAAGCCCAGAAGAAAGCCGAGGCGGAGCGGAAAGCAGAGA  8787
  V  T  E  V  G  L  E  A  Q  K  K  A  E  A  E  R  K  A  E  R  516

GACAAGCCAAGGATTTGGAGAAGGCCCAAAAGTCCTAAACATGACCTGTCGGGTCATCAA  8727
  Q  A  K  D  L  E  K  A  Q  K  S  *                          527

TTTGAACCACAGTCACCGCGTGCTTGCGTTACGAAGCGACAATGATAGACCGATAGACTA  8667
TACGATACCCGTAGCATGAAGCAATTAATGGTGACTTTTACTGTCTAACCATCGTTGCCC  8607
GCCCACACTGATCGTTGGCTTAGGATCCTTCGAGGACAGGAGCAGTCATATTATTAAGTG  8547
CTATGGTCTTGATAATTGTAATGTTGTGATGGACAGTGAAGTGGCATGAGGTGGATGCAA  8487
```

Figure 3.11 Alignment of the Deduced Amino Acid Sequence of *dtp1* with Other Efflux Pumps.

Sequence alignment of the deduced amino acid sequence of the *D. pini dtp1* gene (Dpdte1, Figure 3.10) with that of the *C. carbonum toxA* gene (CcToxA), and the two *S. pombe* transmembrane proteins (SpC3H1 and SpSPBC) is shown. See Table 3.1 for references. Sequence alignment was performed with the GCG PILEUP program using the default settings. All nonpolar hydrophobic residues (L, I, V, M, F, Y, and W) are shaded blue. The possible ATP-binding site is indicated by a cyan-coloured line.


Dpdtp1
CcToxA
SpC3H1
SpSPBC

The *dtp1* ORF is 1692 bp in length (1584 bp without intron sequences, Figure 3.10), encoding a predicted polypeptide product of 527 amino acid residues, which is comparable with the size of the *C. carbonum toxA* gene product but smaller than the two *S. pombe* proteins. The translation initiation codon (ATG) was determined to be at nucleotide position 10440 and was chosen as it was followed by the longest ORF and because the next in-frame start codon was 201 nucleotides downstream, after a region with considerable homology. The putative termination codon (TAA) is at 8747. The nucleotide sequence and deduced amino acid sequence for *dtp1* is shown in Figure 3.10.

Sequence analysis identified two putative introns in *dtp1* (Figure 3.10). The positions of which are not conserved between any of the above genes.

Comparison of the *dtp1* sequence with those of other genes suggest that *dtp1* encodes a efflux pump which belongs to the major facilitator superfamily of transmembrane proteins. Members of this class are highly hydrophobic and typically have 12 membrane spanning domains. The deduced amino acid sequence of dtp1 is also highly hydrophobic, with 217 (41.2%) nonpolar hydrophobic residues (L, I, V, M, F, Y, and W), which is very similar to TOXA (40.7%). The deduced amino acid sequence of *dtp1* was aligned with that of *toxA* and the two *S. pombe* genes. Figure 3.11 shows the alignment with all hydrophobic residues shaded blue, this illustrates the similarity of the gene products with 109 conserved hydrophobic residue sites between all four sequences. As well as conserved hydrophobic regions, a motif (DTP1^{158}FXINLP^{163}) is observed in all four sequences which is also present in many other membrane pumps (Pitkin *et al.*, 1996). It is not clear what the functional significance of this motif is but it could be a possible ATP-binding site.

3.8 *DDH1* DISPLAYS SIMILARITIES TO DEHYDROGENASES.

The *ddh1* (dothistromin dehydrogenase) gene displays similarities to many bacterial polysaccharide dehydrogenase genes. The deduced amino acid sequence of *ddh1* displays 21% identity (38% similarity) to that of the *Pseudomomas aeruginosa wbpA* gene, and 18% (32%) to both the *Staphylococcus aureus cap5O* gene and the UDP-N-acetyl-D-mannosaminuronic acid dehydrogenase (MAD) gene from *Methanococcus jannaschii* (functions and references for these genes are shown in Table 3.1).

Figure 3.12 Nucleotide and Deduced Amino Acid Sequence of the *Dothistroma pini ddh1* Gene.

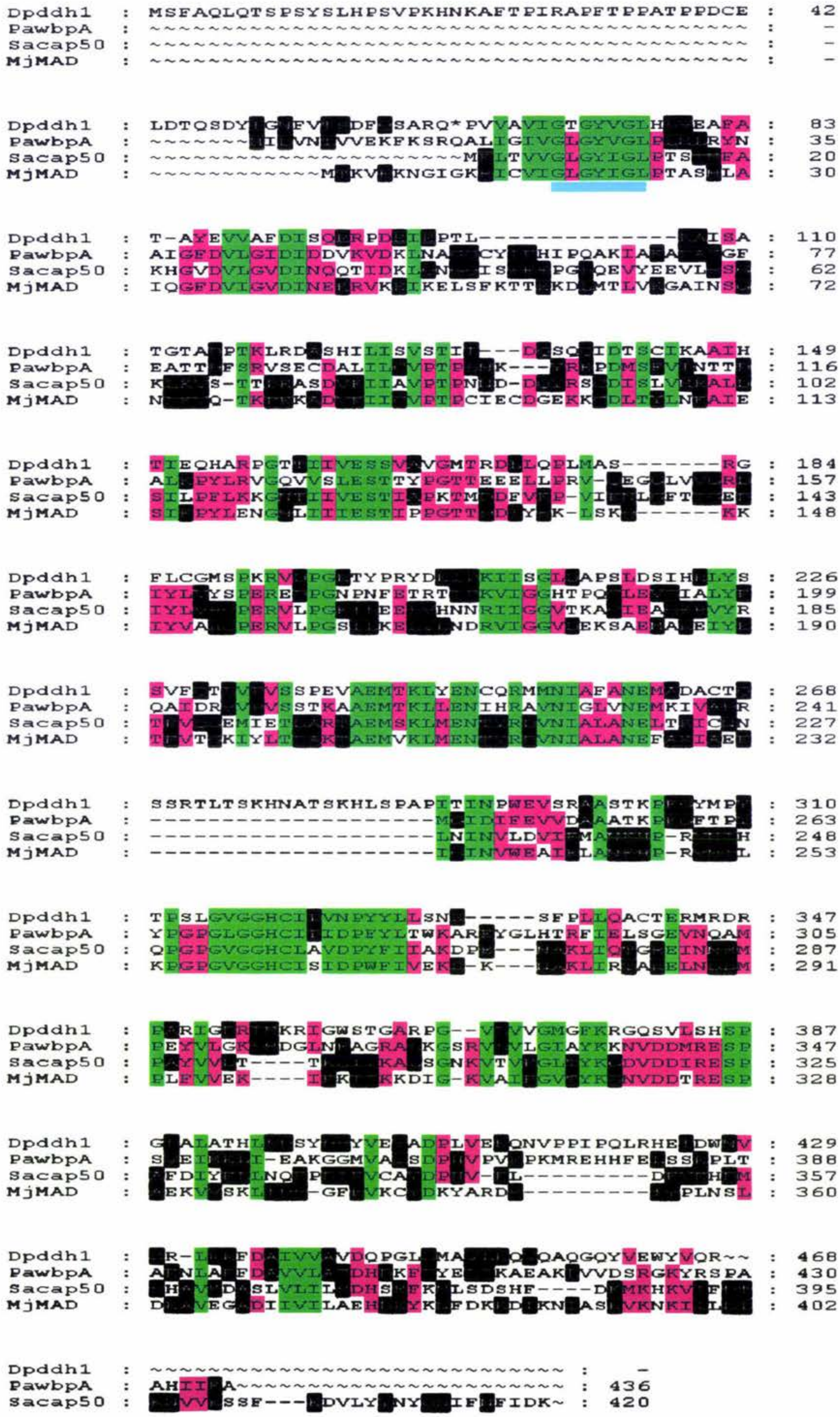
The coding sequence, promoter and terminator regions and deduced amino acid sequence of *ddh1* are shown. Numbers on the right refer to the positions of nucleotides, relative to the start of the λCGV1 sequence (top), and amino acid residues (bottom, red). Initiation and termination codons are shown in green. The premature stop codon is underlined and shown in green. Locations of primers used in RT-PCR are indicated. TATA and CAAT promoter motifs are shown in blue. The 3` end of *dkr1* is indicated.

```
GACGACCATGCCTTGATCTAGCAGTCTGGGTCGTGAGAACTCGCGTCACAATTTAATGAG  1140
ACCTGCCAAATAGCGCCATATGATGCCGGGATCACTCGACCTTGAGTCGTCACAGACGGA  1200
TCTGGTCCGCGGGCATTCACTTGGACGCACAGCCTGCTTACACGAGTCACCGCGCACTCC  1260
GACATGGCTTGAAGCTTCTCGTCGCGTGGCAGCCTGTGCACTGTCGGCGGGGGCGAATGT  1320
GGCTCTCACCATGTCCGCCGGTCGAGATGATCCCTCGCTAGAATTCCCAGAACAGTAAGA  1380
GCCGTCTTGTCCGTAGGCGAGGAGGCGTCGCTCTGTGACATCGCACGCAAGACTATCTGC  1440
CAAGGTCTCTTGCTGATCTATCGCATCTGGACCTTGCCTGATTTGCAACACGTGCTTCAT  1500
CCTATGAAGCTTCCTGAGCGTCCTTGGCCACGGTGTGAACCAAGACATATTGTCCAGACT  1560
ACGATAGAGTGCACAATACAGAGCCACTGGTGCCAAGGTTGCCATGTGCGCCGTTGCGTA  1620
TGGGTCTAGTGTCGAGGCGCTCTTGCGAAAGAACTGCCTACCTCACGCTACCATGTTGGC  1680
TTCCAAGGTAACACGTTGATGGCTTTCGAGGCGGCTGCATACGTCGGGCGCCCCCAACTC  1740
TCATGTGGAAGCCACACTGTGATGGCTTGCTGGGCACAGCGTTGTAAGGCTCTTGTAAAC  1800
ACCAGGCATCAATGGGCGGTATATAAGTCCACACAACGTGAGCCTCATTCTGCTGGCCGT  1860
TCATATTCTGCGGGCAAAGCAACCCTTCAGTCCTATCCTTACTTCACCATCACCAGACCC  1920
TGCCACGATGTCTTTCGCCCAGCTGCAAACCTCTCCTAGCTACAGCCTACACCCGTCAGT  1980
       M  S  F  A  Q  L  Q  T  S  P  S  Y  S  L  H  P  S  V   18

CCCCAAGCACAACAAAGCGTTCACACCCATCAGAGCCCCTTTCACGCCGCCTGCGACACC  2040
 P  K  H  N  K  A  F  T  P  I  R  A  P  F  T  P  P  A  T  P   38

TCCTGACTGCGAGCTCGACACTCAAAGCGATTACATTGGCAACTTTGTGACCTCTGACTT  2100
 P  D  C  E  L  D  T  Q  S  D  Y  I  G  N  F  V  T  S  D  F   58

TCAATCAGCCCGACAATGACCCGTGGTGGCCGTCATTGGCACTGGATATGTCGGCCTTCA  2160
 Q  S  A  R  Q  *  P  V  V  A  V  I  G  T  G  Y  V  G  L  H   78

TCTAGTTGAAGCATTTGCAACAGCATATGAAGTTGTCGCCTTCGACATATCTCAGCGGAG  2220
 L  V  E  A  F  A  T  A  Y  E  V  V  A  F  D  I  S  Q  R  R   98

ACCGGATGAGATCGAGCCTACCCTCCGCGCGATCTCTGCCACGGGGACTGCAGACCCAAC  2280
 P  D  E  I  E  P  T  L  R  A  I  S  A  T  G  T  A  D  P  T   118

GAAGCTTCGAGATGCTTCTCACATCCTGATCTCGGTCAGCACCATCATCGACCAGAGCCA  2340
 K  L  R  D  A  S  H  I  L  I  S  V  S  T  I  I  D  Q  S  Q   138

ACAGATCGACACATCATGCATCAAGGCTGCCATACACACGATCGAGCAACACGCTCGCCC  2400
 Q  I  D  T  S  C  I  K  A  A  I  H  T  I  E  Q  H  A  R  P   158

GGGTACAACGATCATTGTCGAGAGCTCCGTAGCCGTCGGCATGACTAGAGACTTGCTGCA  2460
 G  T  T  I  I  V  E  S  S  V  A  V  G  M  T  R  D  L  L  Q   178

ACCCCTGATGGCTTCCAGAGGCTTCCTCTGCGGCATGTCGCCCAAGCGCGTCGACCCAGG  2520
 P  L  M  A  S  R  G  F  L  C  G  M  S  P  K  R  V  D  P  G   198

AAGAACATATCCTCGATACGACGAGATCCCCAAGATCATCTCCGGCCTCGACGCTCCCTC  2580
 R  T  Y  P  R  Y  D  E  I  P  K  I  I  S  G  L  D  A  P  S   218

TCTCGACTCGATCCACCGCCTCTACTCCTCAGTCTTCCAAACTCTCGTCCCCGTCTCATC  2640
 L  D  S  I  H  R  L  Y  S  S  V  F  Q  T  L  V  P  V  S  S   238

TCCCGAAGTCGCCGAAATGACAAAACTCTACGAAAACTGCCAGCGCATGATGAACATCGC  2700
 P  E  V  A  E  M  T  K  L  Y  E  N  C  Q  R  M  M  N  I  A   258

CTTCGCAAACGAAATGGCCGACGCCTGCACCCAATCATCTCGAACCTTGACTTCCAAACA  2760
 F  A  N  E  M  A  D  A  C  T  Q  S  S  R  T  L  T  S  K  H   278

CAATGCCACGTCGAAACACCTCTCACCCGCGCCCATCACCATCAATCCCTGGGAAGTCTC  2820
 N  A  T  S  K  H  L  S  P  A  P  I  T  T  I  N  P  W  E  V  S   298

                                        DST7ep7➔
TCGCGCCGCAAGCACCAAGCCATTCGGCTACATGCCCTACACGCCCTCCCTGGGAGTCGG  2880
 R  A  A  S  T  K  P  F  G  Y  M  P  Y  T  P  S  L  G  V  G   318

AGGCCACTGCATCCCCGTCAACCCCTACTACCTCCTCTCCAACTCCAGCTTCCCCCTCCT  2940
 G  H  C  I  P  V  N  P  Y  Y  L  L  S  N  S  S  F  P  L  L   338

CCAAGCCTGCACCGAGCGGATGCGGGATCGTCCGGCACGGATAGGAGACCGCATCATGAA  3000
 Q  A  C  T  E  R  M  R  D  R  P  A  R  I  G  D  R  I  M  K   358

GCGAATCGGCTGGTCCACTGGGGCGCGTCCGGGAGTCCTCGTCGTGGGGATGGGCTTCAA  3060
 R  I  G  W  S  T  G  A  R  P  G  V  L  V  V  G  M  G  F  K   378

GAGAGGACAATCAGTGCTCTCGCACTCTCCCGGCCTGGCTCTTGCGACACATTTACTGGA  3120
 R  G  Q  S  V  L  S  H  S  P  G  L  A  L  A  T  H  L  L  D   398

CTCGTATGATGTGTATGTCGAGTATGCGGATCCGTTGGTGGAGGAGCAGAATGTTCCGCC  3180
```


←DS150ep1
GATTCCGCAGCTTCGTCATGAAATTGATTGGAACGTTTCGCGTTTGAGACGGTTTGATGC 3240
I P Q L R H E I D W N V S R L R R F D A 438
GATTGTGGTGGCGGTGGATCAGCCGGGGCTTGATATGGCTGTCTTGGAACAGGTACAGGC 3300
I V V A V D Q P G L D M A V L E Q V Q A 458
TCAAGGGCAGTATGTAGAGTGGTATGTGCAGAGATGAGGTTCATATGGGGCATGCTGGAA 3360
Q G Q Y V E W Y V Q R * 469
GACGATCCTCGCCGTCATGGAGTAATAGCTTGTTTGCATCGGTAGTAAATGAGATATGTC 3420
ACTAGACATCACATCGACCATAACACTCCACTCATGGAGGATAGATACCCGGTGCGTTCT 3480
ATCATGGAATCTAGTGTCCGCTAGATGCCAACAGCAAAAGAAACTTACTTCTTCTCCAGT 3540
TCCCAGTTCAAAGACTTCACACCTATCGGAAAGCACCACCGTCAATACCAAGAATCTTCC 3600
* R F A G G D I G L I K
←dkr1

Figure 3.13 Alignment of the Deduced Amino Acid Sequence of *ddh1* with Other Polysaccharide Dehydrogenases.

Sequence alignment of the deduced amino acid sequence of the *D. pini ddh1* gene (Dpddh1, Figure 3.12) with that of the *P. aeruginosa wbpA* gene (PaWbpA), the *S. aureus cap5O* gene (SaCap5O) and the *M. jannaschii* UDP-N-acetyl-D-mannosaminuronic acid dehydrogenase (MAD) gene (MjMAD) is shown. See Table 3.1 for references. Sequence alignment was performed with the GCG PILEUP program using the default settings. The percentage similarity of *ddh1* to these sequences is nearly twice the percentage identity (see Table 3.1) and this is illustrated by shading: green shading indicates a particular residue (or a similar residue) is present in all four sequences, magenta- three of the four sequences and black shading for similar residues in two of the four sequences. Magneta line indicates the putative NAD-binding site and the cyan line indicates the possible active site.


Dpddh1
PawbpA
Sacap50
MjMAD

The *ddh1* ORF is 1410 bp in length, with the predicted translational initiation codon at nucleotide position 1928 and the termination codon at 3335, and is predicted to encode a product of 469 amino acids. However, a premature stop codon is detected at position 2117. There is no evidence to suggest that this is a sequencing error. If this sequence is correct, the expected product would be a truncated, and probably nonfunctional, polypeptide of 63 amino acids. No introns are predicted in the nucleotide sequence. The nucleotide sequence and deduced amino acid sequence of *ddh1* is shown in Figure 3.12.

Ignoring the premature stop codon, deduced amino acid sequences of *ddh1* and the above genes were aligned (Figure 3.13). A region that is conserved in all the sequences displays the typical NAD binding site consensus motif (^{70}GXGXXG^{75}) expected in dehydrogenases of this type (Burrows *et al.*, 1996). A second highly conserved region is observed (^{316}GVGGHCIPVNPYYL^{329}), the function which is not known although it could possibly be the active site.

3.9 UNTRANSLATED REGIONS OF THE FIVE DOTHISTROMIN GENES.

3.9.1 Promoter Analysis.

The 5` untranslated region of each *D. pini* gene described above was searched for TATA and CAAT motifs which are thought to be important for transcription initiation in filamentous fungi (Hamer and Timberlake, 1987). However, what exact role these factors mediate is uncertain and filamentous fungi are known to be variable with respect to transcriptional initiation motifs. The *D. pini dkr1* gene promoter contains no TATA or CAAT motifs. The *dox1 ddh1*, and *dte1* genes have TATA boxes at 81, 108 and 88 bp (or 153 bp, depending which ATG codon), respectively, upstream of the proposed translational start site. The *dtp1* promoter contains both TATA and CAAT motifs (325 bp and 255 bp upstream from the translatioal start codon, respectively), however the positions of these motifs are both further away from the translation start codon than usual. The constitutively expressed *D. pini* β-tubulin *tub1* gene lacks any of these promoter motifs.

One striking feature of the *D. pini dkr1* gene promoter is a 56 bp region, immediately upstream from the translation initiation codon, which is 84% AC rich, containing 13 AC

(purine/pyrimidine) repeats. This region was inspected in the other reductase genes and found to vary; the *ver1* region is 72% AC rich over 36 bp, *stcU* is 61% AC rich over 31 bp, *brn1* is 70% AC rich over 53 bp and thr1 is 60% AC rich over 30 bp. The values for other dothistromin genes were also determined and also varied; the *dtp1* region is 41% AC rich over 44 bp, *dox1* is 71% AC rich over 55 bp, *dte1* is 64% AC rich over 47 bp and *ddh1* is 70% AC rich over 53 bp. An equivalent region in the *D. pini* β-tubulin gene, *tub1*, promoter is 79% AC rich (over 48 bp). It is not known whether these sequences have any role in either transcription or translation, or if they are involved in the targeting or stability of the mRNA.

AFLR is involved in the regulation of ST/AF biosynthesis and is a GAL4-like transcription factor (Section 1.6.3). The DNA binding site upstream of genes regulated by GAL4 and similar transcription factors PUT3 and PPR1, are two rotational symmetric 3 bp sites (CGG...GCC) separated by 11, 16, and 6 bp respectively (Woloshuk *et. al.*, 1994 and references therein). AFLR has been shown to bind specifically to a palindromic sequence (TTAggccTAA) within its own promoter (Woloshuk *et. al.*, 1994). However, this sequence is not present in all pathway genes and AFLR binding to other cluster gene promoters has not been reported. A search for these sequence types in the promoter regions of the putative dothistromin genes revealed similar motifs but none of these were consistent between genes.

3.9.2 3` Untranslated Region

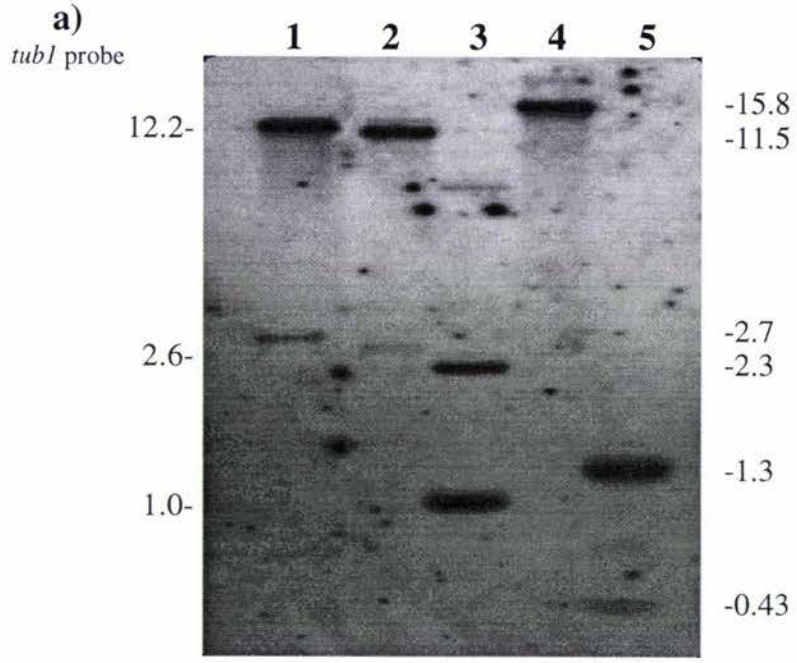
No polyadenylation signal sequences (AATAAA; Proudfoot and Brownlee, 1976) were observed in the 3` untranslated region of the dothistromin genes, except for *dtp1* which has an AATTAA sequence located 104 bp from the termination codon and has tentatively been assigned.

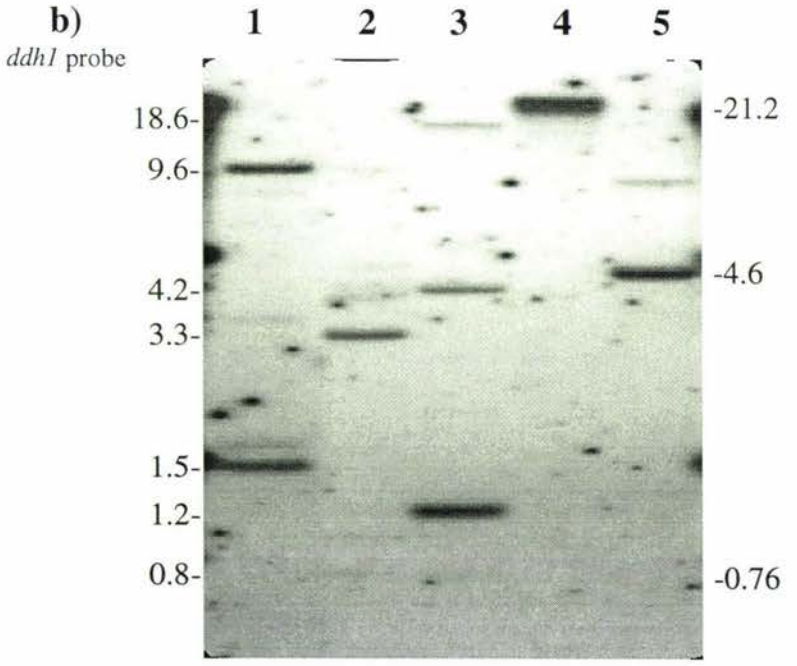
3.10 GENOMIC ORGANISATION OF THE DOTHISTROMIN GENES.

Southern blot analysis of *D. pini* genomic DNA was performed to verify that the λCGV1 clone truly represents the genomic organisation and is not the result of cloning artifacts procured in the construction of the *D. pini* genomic library. This analysis also indicated the copy number of each gene.

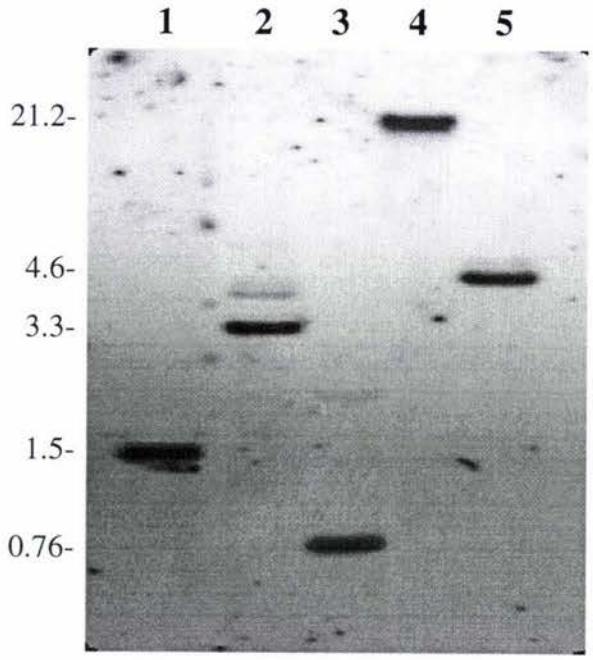
Figure 3.14 *Dothistroma pini* genomic Southern Blots.

Each blot contained 4 µg *D. pini* genomic DNA digested separately with *Bam*HI **(lane 1)**, *Hin*dIII **(lane 2)**, *Sal*I **(lane 3)**, *Xba*I **(lane 4)** and *Xho*I**(lane 5)** and were hybridised with **a)** *tub1* probe- λBT-1 *Eco*RI 1.45 kb. **b)** *ddh1* probe- *Eco*RI 2.7 kb. **c)** *dkr1* probe- *Sal*I 0.76 kb. **d)** *dox1* probe- *Eco*RI/*Bam*HI 3.2 kb. **e)** *dtp1* probe- *Bam*HI 1.7 kb and **f)** *dte1* probe- *Eco*RI 2.1 kb.

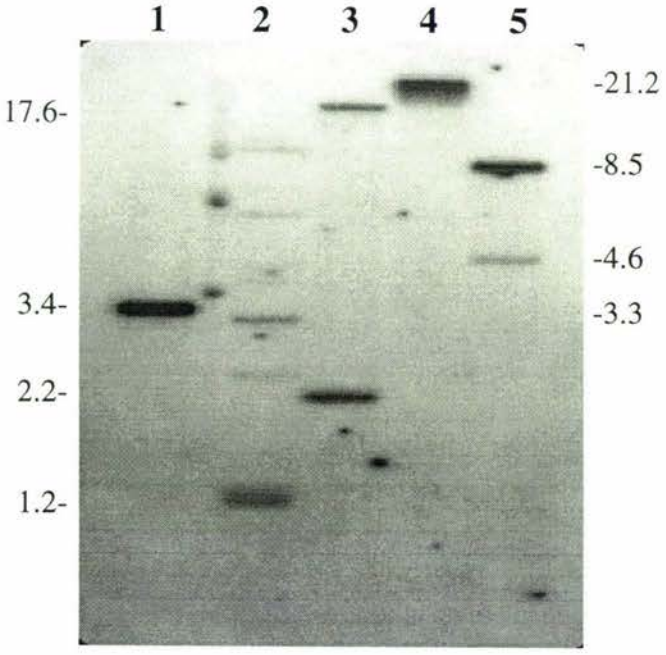

a)
tub1 probe
1
2
3
4
5
12.2-
2.6-
1.0-
-15.8
-11.5
-2.7
-2.3
-1.3
-0.43

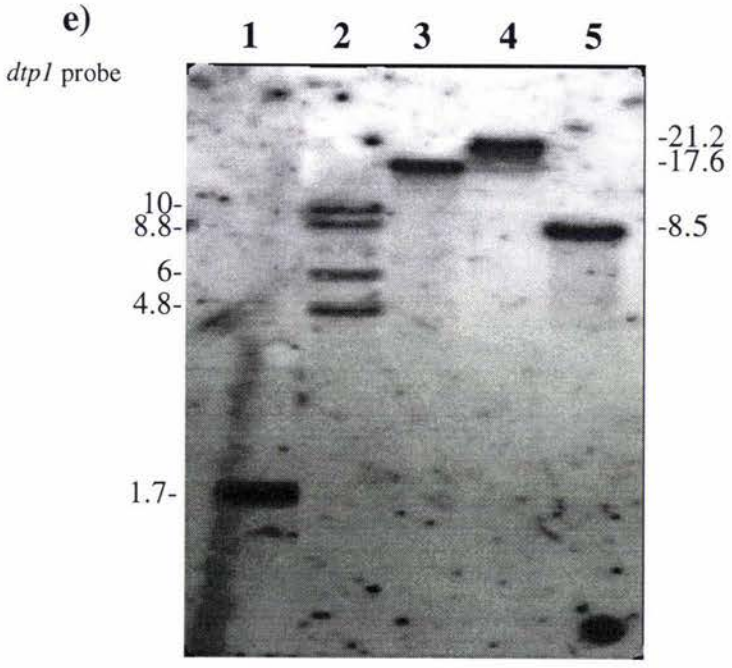

b)
ddh1 probe
1
2
3
4
5
18.6-
9.6-
4.2-
3.3-
1.5-
1.2-
0.8-
-21.2
-4.6
-0.76

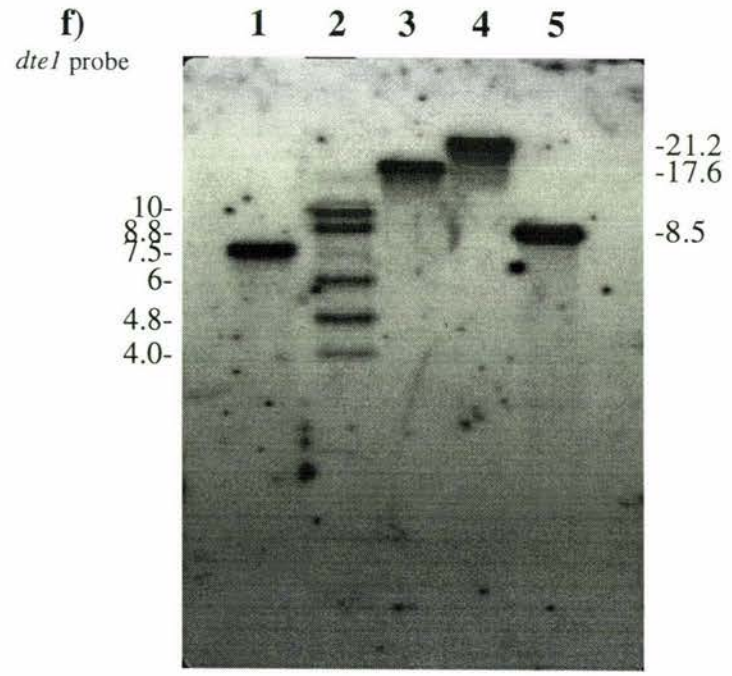
c)
dkr1 probe



d)
dox1 probe




e)
1 2 3 4 5
dtp1 probe
-21.2
-17.6
10-
8.8-
-8.5
6-
4.8-
1.7-


f)
1 2 3 4 5
dte1 probe
-21.2
-17.6
10-
8.8-
7.5-
-8.5
6-
4.8-
4.0-

D. pini ddh1, *dkr1*, *dox1*, *dtp1* and *dte1* gene-specific fragments (Figure 3.3, Figure 3.1) were used as probes and hybridised to Southern blots containing 4 µg *D. pini* genomic DNA digested separately with *Bam*HI, *Hin*dIII, *Sal*I, *Xba*I and *Xho*I. Two identical blots were produced.

[α-^{32}P]dCTP-labelled probes hybridised to blots were: *tub1* probe (positive control containing a section of the *D. pini* β-tubulin gene)- λBT-1 *Eco*RI 1.45 kb (Figure 3.14a); *ddh1* probe- *Eco*RI 2.7 kb (Figure 3.14b); *dkr1* probe- *Sal*I 0.76 kb (Figure 3.14c); *dox*1 probe- *Eco*RI/*Bam*HI 3.2 kb (Figure 3.14d); dtp1 probe- *Bam*HI 1.7 kb (Figure 3.14e) and *dte1* probe- *Eco*RI 2.1 kb (Figure 3.14f). Hybridisations were performed at 65°C and the blots were washed at 65°C with 0.5 x SSC and 0.1% SDS.

Unfortunately, some partial fragments were observed in the *Hind*III and *Sal*I lanes, however, the fragment sizes of these bands concurred with the expected sizes for partial digestion (Figure 3.14, Figure 3.1). Explanations of these questionable bands follow.

The *ddh1* probe cross-hybridised to the λ size markers, thus effectively titrating the probe and reducing hybridisation to the genomic *D. pini* region (Figure 3.14b). This resulted in very faint bands, especially for the small fragments. To increase intensity, the blots were put down for autoradiography for much longer than usual. Unfortunately, small bands were still faint and partial digest bands were observed which would not usually have been, however, all bands could be accounted for (for example the *Sal*I 4.2 kb band is a partial fragment composed of *Sal*I 1.2, 0.76 and 2.2 kb bands).

The *dte1* probe was expected to hybridise to two *Hin*dIII fragments, one of 4.8 kb in size (as determined by λCGV1 sequence analysis) and the other of unknown size (but greater than 0.6 kb). Instead, the *dte1* probe hybridised to five *Hin*dIII fragments; 10 kb, 8.8 kb, 6.0 kb, 4.8 kb and 4.0 kb. The 4.8 kb band is expected, and because the 4 kb band is obviously smaller than the 4.8 kb band (therefore can not be a result of a partial digest), the 4.0 kb band is assumed to be the second expected band. It is proposed that the 6.0 kb band is a partial digest of the *Sal*I 4.8 and 1.2 kb bands, the 8.8 kb band is a partial digest of the 4.0 kb and 4.8 kb bands and the 10 kb band is a partial digest of the 4.8, 1.2 and 4.0 kb bands (Figure 3.14f and Figure 3.1). The *Hin*dIII lane of the *dtp1* hybridised blot

(Figure 3.14e, lane 2) produces a similar banding pattern to that above, when only a single band should hybridise (4.8 kb). The explanation for these extra bands is the same as that for the *dte1* hybridisation pattern.

Overall, Southern analysis clearly showed that the λCGV1 sequence and organisation is representative of the *D. pini* genome. Results also indicate that all genes are present in single copies; this is especially evident in *Bam*HI digests for each of the hybridisations (Figure 3.14: Lane 1). The restriction enzyme *Xba*I does not cut within the 13.3 kb genomic region sequenced, which is illustrated by the same band (~21.2 kb) being present in each of the λCGV1 gene hybridisations. Because no bands other than the 21.2 kb band are observed, this confirms that the genes are clustered.

CHAPTER THREE- DISCUSSION.

3.11 IDENTIFICATION OF A PUTATIVE DOTHISTROMIN CLUSTER.

Little is known about the mechanics of dothistromin biosynthesis. The chemical and biochemical data available suggests a close biogenetic relationship between dothistromin and AF biosynthesis (Section 1.4). In investigating dothistromin biosynthesis at the molecular level, we are not in a position of isolating specific genes for particular steps, but identifying likely dothistromin biosynthetic genes based on similarities to ST/AF biosynthesis and elucidating the function of each isolated gene by gene disruption.

The λCGV1 clone was isolated from a *D. pini* genomic library based on sequence similarity to the *A. parasiticus ver1* gene (Gillman, 1996). In this study, λCGV1 was completely sequenced and analysis of the 13.3 kb *D. pini* genomic region identified five genes. Based on similarities to AF genes, these genes are thought to be involved in dothistromin biosynthesis and have collectively been tentatively termed dothistromin genes. Southern analysis confirmed the genomic clustering of the dothistromin genes (Section 3.10; Figure 3.14).

3.11.1 AF-Like Genes are Clustered in *D. pini.*

Three of the proposed dothistromin genes (*dkr1*, *dox1* and *dte1*) show significant similarities to genes involved in ST/AF biosynthesis.

3.11.1.1 *dkr1*

The *dkr1* gene displays high amino acid sequence identity to the *ver1* and *stcU* genes that are required for ST/AF production (Table 3.1; Section 3.4). The *ver1* and *stcU* genes each encode a NADPH-dependent ketoreductase involved in the meta-diphenol dehydroxylation of versicolorin A to ST (Skory *et al.*, 1992; Keller *et al.*, 1994; Figure 1.2). DKR1 also displays strong similarities to genes involved in polyphenol dehydroxylations in melanin biosynthesis (Vidal-Cros *et al.*, 1994; Perpetua *et al.*, 1996; Shimizu *et al.*, 1997; Table 3.1). Between all genes there is complete conservation of a NADPH-binding motif and the active site (Figure 3.5). The high sequence identity of *dkr1* with the genes presented here indicates that *dkr1* is a NADPH-dependent ketoreductase involved in a meta-diphenol dehydroxylation step in dothistromin biosynthesis. To confirm this and identify the substrate of *dkr1*, recombinational inactivation of *dkr1* needs to be performed.

Interesting questions are raised due to the exceptionally high similarity of *dkr1* to the genes discussed above. Firstly, if *dkr1* is involved in dothistromin biosynthesis, why such high identity to a gene involved in AF production considering that neither gene (*dkr1* or *ver1*) is required for growth? There is no evidence that AF is involved in pathogenesis, as there is for dothistromin, therefore there seem to be different roles for the toxins. This may suggest that the genes are evolutionarly linked and horizontal transfer of the sequences between fungi may have taken place. This is investigated in Chapter 5. There are no reports that *D. pini* can produce aflatoxins, and because *dkr1* has been shown to be expressed (Figure 4.2), it is unlikely that the *dkr1* sequence is non-functional or alternatively a functional copy of a versicolorin A reductase gene involved in AF biosynthesis. Recent research in our laboratory has identified overseas *D. pini* isolates which produce very high levels of dothistromin (compared to the New Zealand Dp2 isolate), these cultures should be tested for AFB_1 to investigate this possibility. Whether dothistromin is a byproduct of AF production should also be investigated.

Whether *dkr1* is a functional homolog of *ver1* and *stcU* genes is an interesting question for future research. This would highlight the relationship between the two pathways. Such experiments may also provide insights into the production and programming of polyketides; whether polyketide pathways are linked and able to shunt intermediates into

other pathways and produce other toxic products. The melanin biosynthetic genes of *Alternaria alternata* restored melanin production in *Magnaporthe grisea* mutants even though the genes did not cross-hybridise in Southern blots, indicating low sequence identity between the biosynthetic genes (Kawamura *et al.*, 1997).

High stringency Southern blot analysis in this study indicated a single copy of *dkr1* was present in the *D. pini* genome (Section 3.10; Figure 3.14c). However, low stringency blots with the *A. parasiticus ver1* probe indicated a second copy of a *ver*-like sequence (Gillman, 1996). This other *ver*-like region has recently been sequenced in our laboratory and results suggest that this gene may be involved in melanin biosynthesis. The deduced amino acid sequence of this gene, *ver2*, exhibits 68% identity to that of *dkr1* (Bradshaw, personal communication). It is not known whether *ver2* is expressed or functional.

3.11.1.2 dox1.

The *D. pini dox1* gene displays a moderate level of identity to *stcC* (Table 3.1), an *A. nidulans* ST cluster gene. The role of *stcC* in ST biosynthesis is uncertain but *stcC* transcripts are observed in the coordinate and specific manner shown for ST cluster gene expression. Both *stcC* and *dox1* show similarities to the well characterised chloroperoxidase gene of *Caldariomyces fumago*. The main region of identity between the amino acid sequences of the three genes is the heme-binding site (DOX1 ^{28}PCPALNALANHG^{39}; Figure 3.7).

The *Caldariomyces fumago* chloroperoxidase enzyme catalyses halogenation reactions and is involved in the biosynthesis of chlorinated secondary metabolites. The *C. fumago* chloroperoxidase protein has received much attention for its level of secretion (500mg/litre; Axley *et al.*, 1986) and as an extremely versatile catalyst, with similarities to peroxidaes, catalases and P-450 cytochromes. (Blanke and Hager, 1988, 1990 and references therein). Studies determined that the chloroperoxidase residues C^{29} and H^{38} (C^{50} and H^{59} in Figure 3.7 because signal peptide is present) as critical residues in the active site which are thought to be important in the versatility of the enzyme (Blanke and Hager, 1988, 1990). These residues are conserved in the *dox1* deduced amino acid sequence.

What role *dox1* has in dothistromin production is not known, and like *dkr1*, gene disruption of *dox1* is required to help elucidate the function of DOX1. Sequence similarities indicate that *dox1* encodes a heme-binding oxidase. Whether DOX1 has a broad specificity, like chloroperoxidase, and is involved in many steps or reactions is unknown but an interesting focal point for future investigation of this gene.

3.11.1.3 *dte1*.

The *D. pini dte1* gene shows identity to the thioesterase domain of many fungal PKSs, including the *A. nidulans* (STCA) and *A. parasiticus* (PKSA) PKSs involved in ST/AF biosynthesis (Table 3.1). PKSs are generally one of two types: type I PKSs consist of very large multifunctional proteins (encoded for by one gene) in which the catalytic sites are domains, whereas type II PKSs consist of separate proteins (encoded for by separate genes) for each catalytic site (reviewed in Hopwood, 1997). The *dte1* gene shows a distinct ORF and thus appears to encode a distinct monofunctional protein, indicating that it could form part of a type II PKS.

Type II PKSs are present mainly in bacteria and a literature search did not identify any examples of a type II PKS in filamentous fungi. The *dte1* gene is situated near the end of λCGV1 (Figure 3.3). To confirm that *dte1* is an independent thioesterase enzyme, and not the start of a type I PKS, the continuing genomic sequence needs to be obtained. It would be unusual though to find a thioesterase domain at the N-terminus of a type I PKS since it is characteristically found at the C-terminus.

A λ clone containing a PKS-like sequence was recently isolated from a *D. pini* genomic library in our laboratory (Morgan, 1998). Sequencing of a 2.4 kb region identified two putative PKS domains (the β-ketoacyl synthase domain and the acyl transferase domain), both of which display strong similarities to the corresponding domains in STCA and PKSA. These domains are encoded by a single ORF, indicating a likely type I PKS. Southern analysis suggests that a single copy of *dte1* is present within the *D. pini* genome (Figure 3.14), intimating that the putative type I PKS may not contain a thioesterase domain. Linkage between the dothistromin genes with this putative type I *pks* gene has not yet been shown.

Thioesterase releases the polyketide product from the enzyme complex by hydrolysing the bond between the completed polyketide and the 4`-phosphopantetheine prosthetic group on the acyl carrier domain of the last chain-extending module. The exact role of *dte1* in dothistromin biosynthesis is not known and gene disruption of *dte1* is required to elucidate function. It will be interesting to determine if *dte1* is specific for dothistromin biosynthesis.

3.11.2 Other Putative Dothistromin Genes.

3.11.2.1 *dtp1*

The *D. pini dtp1* gene displays similarities to many small efflux pumps belonging to the drug resistance protein family of the major facilitator superfamily of transmembrane proteins (Table 3.1). The deduced amino acid sequence of *dtp1* displays 24% identity (41% similarity) to the *Cochliobolus carbonum* HC-toxin efflux pump encoded by *toxA* (Pitkin *et al.*, 1996).

HC-toxin is a cyclic tetrapeptide, produced by a peptide synthetase encoded by *hts1*, which is required for the pathogenicity of *C. carbonum* on maize (Panaccione *et al.*, 1992). HC-toxin is secreted from the fungus and there is some evidence that HC-toxin is toxic to *C. carbonum* (Pitkin *et al.*, 1996). The *toxA* and *hts1* genes are divergently transcribed with the two transcriptional start sites 386 bp apart. This proximity may be necessary for coordinate regulation of *toxA* and *hts1*, ensuring efficient export of HC-toxin from the cell, thus protecting the cell. This arrangement is common in bacteria (Pitkin *et al.*, 1996).

In culture, *D. pini* secretes dothistromin from the cell, producing dark brown/red areas which surround the colony. Dothistromin has been shown to inhibit RNA synthesis in *Chlorella pyrenoidosa* and *Bacillus megaterium*, indicating that dothistromin may be toxic to *D. pini* itself and export of dothistromin from the cell may also be a protection mechanism (Harvey et al., 1976). This implies that *D. pini* has a dothistromin efflux pump.

Based on the similarities to the proteins described and the proximity to likely dothistromin genes, *dtp1* is proposed to encode an efflux pump, that belongs to the major facilitator superfamily of transmembrane proteins, and is involved in the export of dothistromin from the cell.

The presence of *dtp1* in the putative dothistromin biosynthetic cluster is interesting. Whether proximity to biosynthetic genes is required for efficient regulation of *dtp1* to ensure proper export of dothistromin, as is thought for the close linkage of *toxA* and *hts1*, remains to be shown. Another possible explanation is that the toxin pump is part of the cluster as a requirement, it is required for the host of the cluster to survive, therefore for the cluster to survive (selfish gene cluster model; discussed in Section 1.7.1).

If dothistromin is toxic to *D. pini*, obtaining a *dtp1* mutant may be difficult as the mutation may be lethal. At present, there are no known conditions which support the growth of the fungus, but not dothistromin production. An effort to determine these conditions is now required. These conditions would allow the propagation of *dtp1* knockout mutants; to test *dtp1* function, the fungus would be switched to dothistromin producing conditions. Determination of these conditions would also allow detailed expression analysis of biosynthetic genes, thus confirming that they are expressed only in dothistromin producing conditions. The lack of a sexual cycle prevents any genetic crosses being done.

3.11.2.2 ddh1.

The *D. pini ddh1* gene displays similarities to many bacterial polysaccharide dehydrogenase genes involved in lipopolysaccharide antigen biosynthesis. DDH1 displays 21% amino acid sequence identity (38% similarity) to the *Pseudomomas aeruginosa wbpA* gene, including conservation of the NAD-binding domain (Burrows *et al.*, 1996). The *wbpA* gene is part of a lipopolysaccharide O-antigen biosynthetic gene cluster in *P. aeruginosa*.

The deduced amino acid sequence of *ddh1* indicates that *ddh1* contains a premature stop codon which would produce a truncated, and probably non-functional, protein. To confirm this, the region should be resequenced.

The *ddh1* sequence displays no similarity to any eukaryotic genes. Whether the presence of *ddh1* is indicative of a bacterial origin for itself, or some of the dothistromin genes, is not known. Evolutionary aspects of the dothistromin genes are discussed in Chapter 5. If the sequence of *ddh1* is incorrect and *ddh1* does encode a functional protein, there is no obvious role for a polysaccharide dehydrogenase in dothistromin biosynthesis. Sequencing regions 5' to *ddh1* may distinguish whether it is a part of the gene cluster or possibly a bordering gene.

3.12 SUMMARY.

A 13.3 kb *D. pini* genomic region was sequenced along both strands and was found to contain five putative genes (Table 3.1; Figure 3.3). Three of these genes, *dkr1*, *dte1*, and *dox1*, are proposed to be involved in dothistromin biosynthesis, as these genes display significant similarities to genes required for the similar ST/AF biosynthetic pathway (Table 3.1). The *dtp1* gene, located between *dox1* and *dte1*, shows strong similarities to members of the major facilitator superfamily of transmembrane proteins and is thought to encode a transmembrane dothistromin efflux pump. The *ddh1* gene may contain a premature stop codon and its role in dothistromin biosynthesis is uncertain. Critical to the explication of the function of each gene, genetic disruption of each gene, especially *dkr1*, is required. The expression of these genes is discussed in Chapter 4 and the evolutionary relationship between dothistromin and ST/AF biosynthesis is investigated in Chapter 5.

4. EXPRESSION ANALYSIS OF PUTATIVE DOTHISTROMIN BIOSYNTHETIC GENES.

4.1 INTRODUCTION.

Sequencing of a 13.3 kb *D. pini* genomic region identified a likely dothistromin biosynthetic cluster which shows similarities to ST/AF gene clusters (Chapter 3). ST/AF cluster genes are coordinately expressed specifically in AF producing conditions. Expression of ST/AF genes correlates with ST/AF production and is observed around the transition from exponential growth to stationary phase onwards (Section 1.6.3).

To investigate the expression of the putative dothistromin genes, RT-PCR was performed using RNA extracted from mycelia samples which were taken in conjunction with a growth curve. Dothistromin production was also monitored over the time course using a competitive ELISA protocol. Analysis identified a non-coordinate pattern of expression with some dothistromin genes being constitutively expressed.

4.2 CONSTRUCTION OF A *D. PINI* GROWTH CURVE.

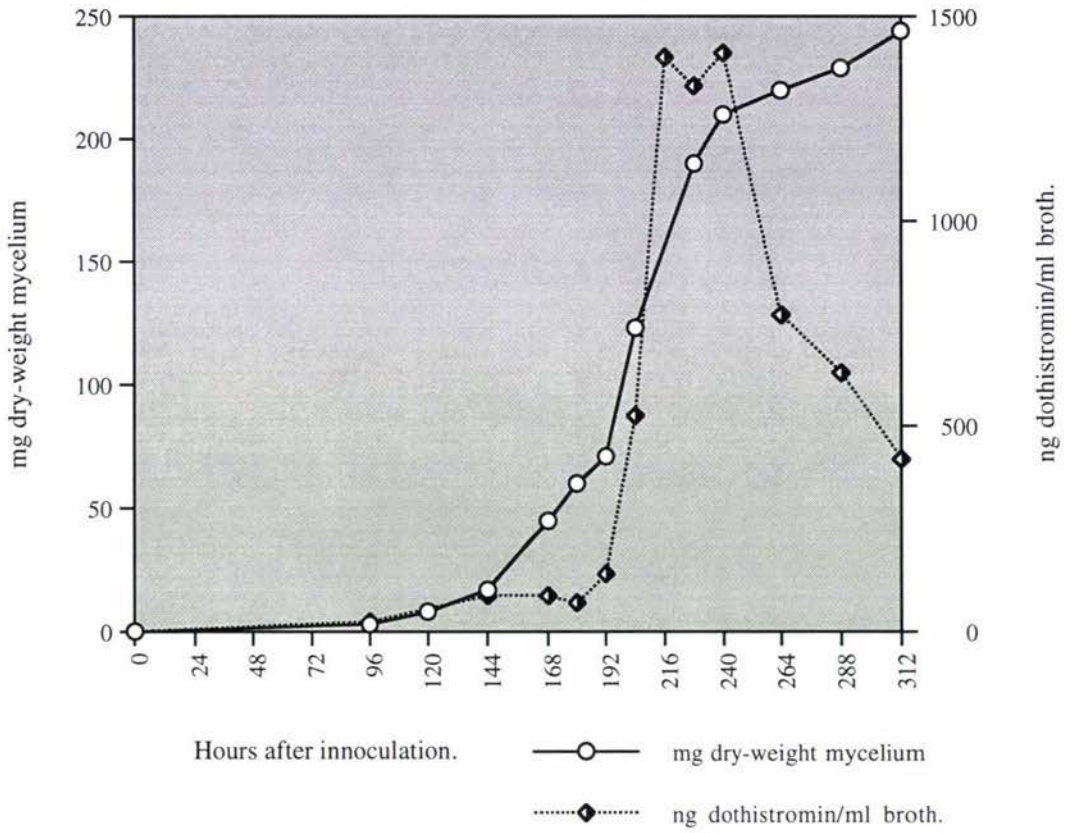
The growth rate of *D. pini* Dp2 was assessed by dry weight determination. Two 2 L flasks containing 800 ml DB (Section 2.2.4) were inoculated with 1 x 10^5 *D. pini* spores per ml of broth and incubated at 20°C (in the dark) with shaking. A total of thirteen 20 ml samples were taken from each flask, as well as a 1 ml broth sample at day zero. Samples from each flask were taken every 24 hours from day four (96 hours) onwards, then every 12 hours between days 7 (168 hours) and 10 (240 hours).

Each 20 ml sample was filtered through miracloth with a 1 ml broth sample being kept for ELISA analysis. Mycelium was washed with water, snap frozen in liquid nitrogen and stored at -70°C until the completion of the experiment. This produced two complete sets of mycelia and broth samples. The wet weight values for both sets of mycelia samples were determined and results indicated that both sets of data were representative of each other. One set was freeze-dried for dry-weight determination while the other set was used for RNA extractions.

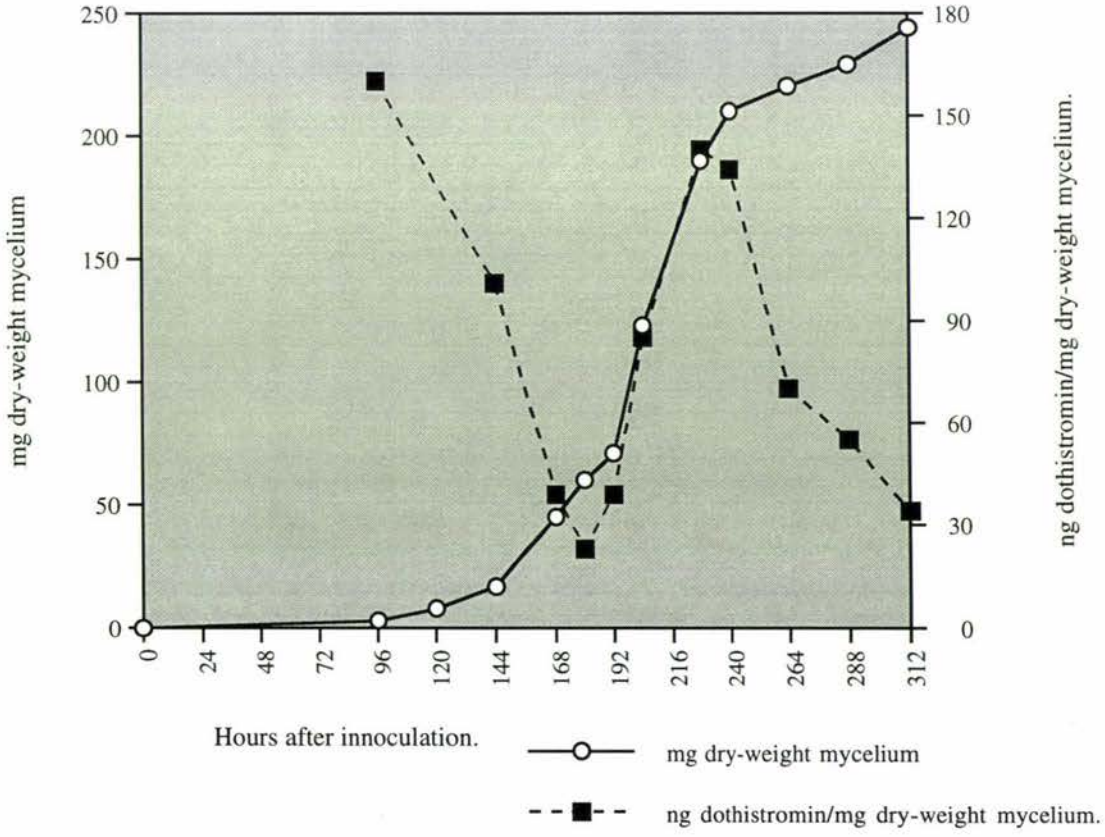
Figure 4.1 Dothistroma pini Growth Curves and Dothistromin Concentration.

Graphs showing growth (mg dry-weight mycelia) of *D. pini* over a 13 day (312 hours) period. Also shown is the dothistromin concentration at the same time points. The two graphs have the dothistromin concentration expressed differently: **A)** ng dothistromin per ml broth and **B)** ng dothistromin per mg dry-weight mycelium.

a)


mg dry-weight mycelium
ng dothistromin/ml broth.
Hours after innoculation.
mg dry-weight mycelium
ng dothistromin/ml broth.

b)


mg dry-weight mycelium
ng dothistromin/mg dry-weight mycelium.
Hours after innoculation.
mg dry-weight mycelium
ng dothistromin/mg dry-weight mycelium.

D. pini is a slow growing fungus in culture, therefore increasing the risk of contamination. Although strict aseptic techniques were adhered to, possible contamination of the *D. pini* cultures (both flasks) was noticed at day 12 and was definitely noticeable at day 13 and the experiment was stopped (24 hours eariler than planned).

The growth curve constructed using mycelium dry-weight produced a standard growth curve with clear lag, exponential and stationary phases and is shown in Figure 4.1.

4.3 DOTHISTROMIN QUANTIFICATION.

To investigate whether production of dothistromin is representative of secondary metabolites, competitive enzyme-linked immunsorbent assays (ELISA) were carried out to determine dothistromin concentrations. As mentioned in Section 3.11.2.1, dothistromin is exported from the cell into the media. Thus the concentration of dothistromin in broth samples is assumed to be representative of true dothistromin production.

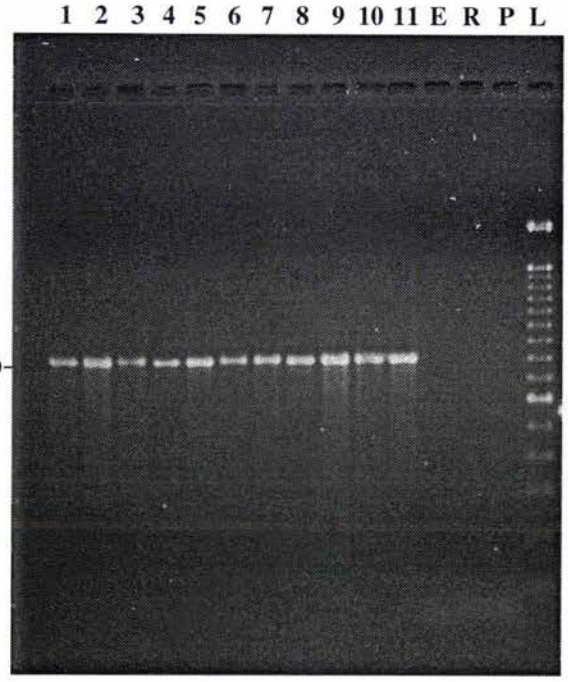
ELISA was performed as in Section 2.18, using the broth samples from the RNA set. Figure 4.1a shows the dothistromin concentrations presented as ng dothistromin per ml of broth. Low concentrations of dothistromin (24-140 ng dothistromin per ml broth) were detected in the early time-course samples (days 5-8, which corresponds to the lag-phase and the start of expotential growth phase). This result is unexpected for normal secondary metabolite production and possible explanations are discussed later in this chapter. However, apart from this result, the resulting trend is indicative of secondary metabolite production, with the rapid appearance of dothistromin occuring late in the exponential growth phase (Figure 4.1a).

However, when the dothistromin concentration is presented as ng dothistromin per mg dry-weight mycelium, a biphasic pattern of dothistromin biosynthesis over growth is observed (Figure 4.1b). The dothistromin concentration pattern for days 8-13 is the same for both graphs, indicating that the different pattern is due to the early dothistromin concentrations (days 5-8).

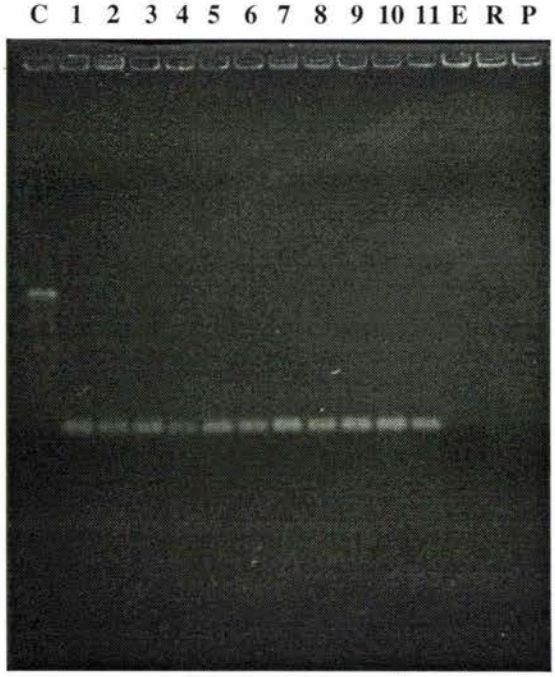
Figure 4.2 Expression Analysis of the Dothistromin Genes.

RT-PCR results are shown. **a)** positive control using the β-tubulin primers tub1 and tub2; **b)** *dte1* primers 151Fep3 and 151Rep2; **c)** *dox1* primers 152Rep5 and 152Fep1; **d)** *dkr1* primers RTFexon1 and RTCexon3; **e)** *dtp1* primers MF4151p3 and MF4152p3; **f)** *ddh1* primers DST7ep7 and DS150ep1. Lane naming is the same for each gel: **lane 1**- day 6 (144 hours) RNA sample; **lane 2**- day 7 (168 hrs); **lane 3**- day 7.5 (180 hrs); **lane 4**- day 8 (192 hrs); **lane 5**- day 8.5 (204 hrs); **lane 6**- day 9 (216 hrs); **lane 7**- day 9.5 (228 hrs); **lane 8**- day 10 (240 hrs); **lane 9**- day 11 (264 hrs); **lane 10**- day 12 (288 hrs); **lane 11**- day 13 (312 hrs); **L**- 100 bp ladder (Life Technologies); **C**- positive control (β-tubulin primers); **E**- negative control (no RT enzyme); **R**- negative control (no RNA); **P**- negative control (no primers). The * symbol in gel c) means that this lane contains the same as lane 1 (a mistake, it was loaded twice instead of the positive control which did work).

a)
tub1 primers



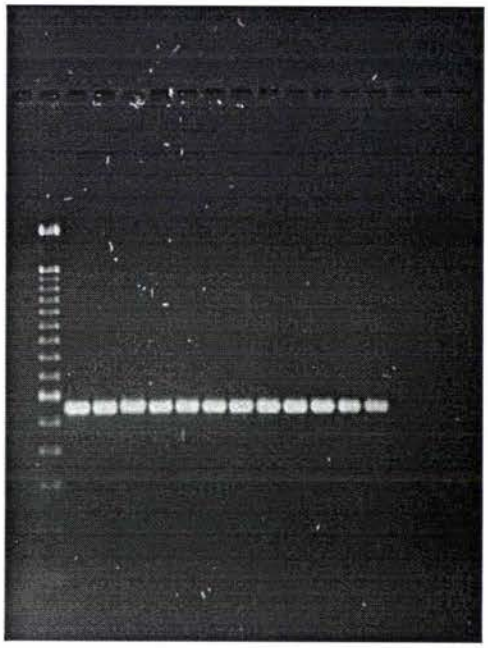
b)
dte1 primers



c)
dox1 primers

L * 1 2 3 4 5 6 7 8 9 1011 E R P

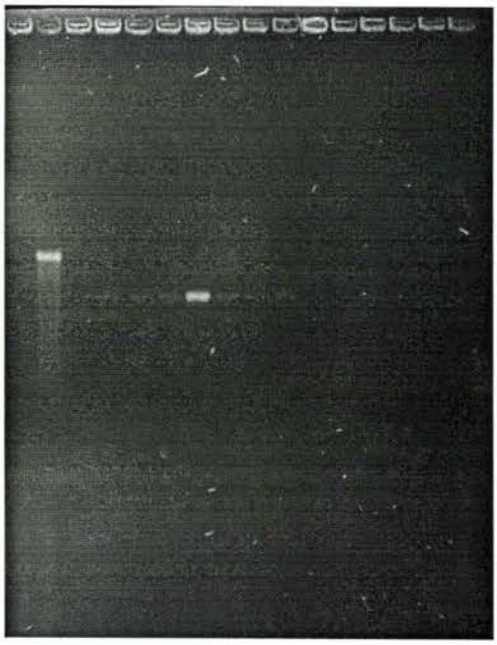
560 bp-



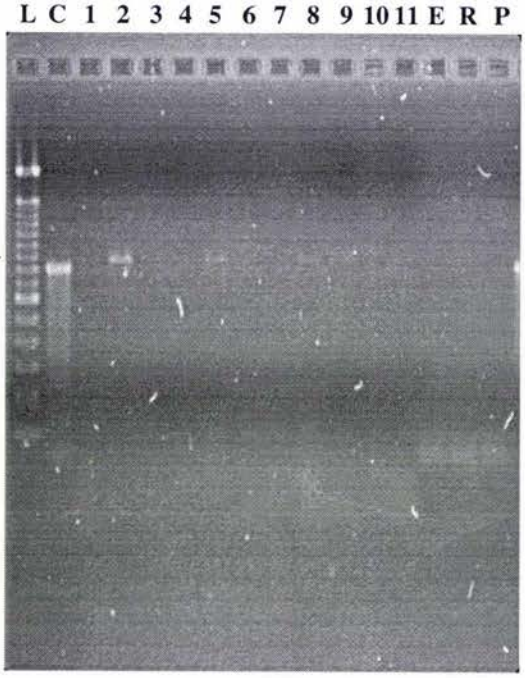
d)
dkr1 primers

C 1 2 3 4 5 6 7 8 9 10 11E R P

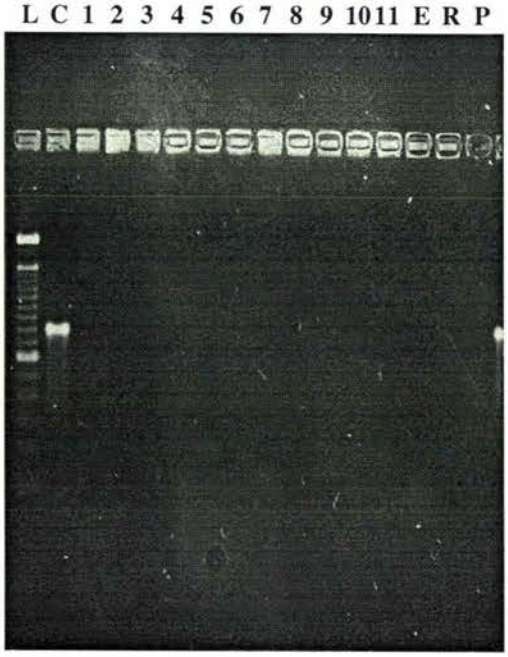
800 bp-
587 bp-



e)
dtp1 primers



f)
ddh1 primers



A decline in dothisromin concentration is observed near the end of the time course (Figure 4.1). This is unexpected, as a stationary level of dothistromin concentration (in ng dothistromin per ml of broth) would be anticipated if *D. pini* had stopped producing dothistromin. A possible reason for this decline, other than experimental error, would be any modification of the dothistromin epitope (the furan ring), preventing monoclonal antibody binding for ELISA detection. The decline in dothistromin concentration coincides with the observation of contamination of the cultures. Whether the contaminant is metabolising dothistromin and thus altering the epitope (therefore altering dothistromin detection) is not known.

4.4 EXPRESSION OF DOTHISTROMIN GENES.

RNA was extracted from the day 6-13 mycelium samples by the TRIzol method (Section 2.15), quantified (Section 2.15.2), DNaseI treated (Section 2.15.2) and stored in 10 µl aliquots (100 ng RNA/µl) at -70°C until use.

4.4.1 RT-PCR.

To determine the presence of specific mRNA transcripts, primers specific to each of the five putative dothistromin genes were used in RT-PCR reactions. These primers were: *dkr1*- RTFexon1 and RTCexon3 (primer positions shown in Figure 3.4), *dox1*- 152Rep5 and 152Fep1 (Figure 3.6), *dtp1*- MF4151p3 and MF4152p3 (Figure 3.10), *dte1*- 151Fep3 and 151Rep2 (Figure 3.8), *ddh1*- DST7ep7 and DS150ep1 (Figure 3.12). Primers (Tub1 and Tub2) to the constitutively expressed *D. pini* β-tubulin gene, *tub1*, were used as a positive control.

RT-PCR reactions were performed using the SuperScript One-Step™ RT-PCR system (Life Technologies; as in Section 2.17). RT-PCR conditions were optimised using the β-tubulin primers (10 pmols of each primer), with varying concentations of total RNA (25 ng- 500 ng), Mg^{2+} (1.5 µM- 3.0 µM; the kit reaction mix has a Mg^{2+} concentration of 1.5 µM) and with two different cDNA synthesis temperatures (50°C and 55°C). The optimal conditions were determined to be: 100 ng total RNA, Mg^{2+} concentration of 1.5 µM and a cDNA synthesis temperature of 50°C. The number of PCR cycles (40) was not altered as the bands produced were bright and clear. Because of limited resources these RT-PCR

conditions were used for each experiment (that is, conditions were not optimised for each different primer pair).

RT-PCR results are shown in Figure 4.2. Two dothistromin genes *dte1* (Figure 4.2b), and *dox1* (Figure 4.2c) exhibited constitutive expression. The product amplified with the *dkr1* primers (587 bp) was clearly visible at day 8.5 (204 hours), which is 12 hours behind the main dothistromin peak, but is weak in the other lanes (Figure 4.2d; analysis was performed three times with the same result each time). The *dtp1* product (907 bp) was faint and not in all of the lanes (Figure 4.2e). The *ddh1* gene showed no obvious expression (Figure 4.2f). Overall, there were a number of different expression patterns, indicating that the genes are not coordinately expressed. Possible explanations for this are discussed later.

To qualitate expression so comparison between lanes could be made, multiplex RT-PCR was attempted using the β-tubulin and *dkr1* primers. This was unsuccessful, as only the β-tubulin product was observed in all lanes (data not shown), indicating that further optimisation of conditions is required.

4.4.2 Northern analysis

Northern blot analysis was attempted prior to RT-PCR being carried out. Although northern blot analysis has many advantages such as it enables quantification, sizing of the transcript, and can be reused, the procedure is not as sensitive as RT-PCR. 150 ml flasks containing 20 ml DB were inoculated with 10^5 spores per ml (similar to above), and were grown at 20°C, with shaking, for 7 and 10 days. Mycelium was harvested and total RNA extracted, quantified, and DNaseI treated (as above). Northern blots were prepared (Section 2.16) containing 8 μg *D. pini* total RNA from each sample. The northern blot was hybridised with the β-tubulin λBT-1 EcoRI 1.45 kb fragment as the positive control. A band of the expected size was observed in both lanes (7 day and 10 day). The blot was then stripped and hybridised with the *dkr1* probe (λCGV1 *Sal*I 0.76 kb). No signals were detected in either lane. Likewise, no signals were observed when hybridised with the *dte1* probe (λCGV1 *Eco*RI 2.1 kb).

CHAPTER FOUR- DISCUSSION.

4.5 DOTHISTROMIN PRODUCTION.

ELISA was used to determine the pattern of dothistromin production. The results were conflicting depending on the manner in which dothistromin production was presented (Figure 4.1). The early time point dothistromin concentrations were thought to be responsible for these differing results, as the pattern observed for the latter time points was the same in both graphs.

The competitive ELISA protocol for dothistromin detection (Jones *et al.*, 1993) was not designed for determination of dothistromin concentrations in growth medium. However, prilimary analysis using known dothistromin concentrations in DB samples indicated that the broth did not interfer with analysis. Further competitive ELISA analysis using broth samples has found the dothistromin concentration to vary considerably between duplicates especially for low dothistromin concentrations (R. Ganley, personal communication). Although this may imply that the low dothistromin values in this study are inaccurate, the results still indicate that dothistromin is present in the early time-point samples. This could be due to: something in the broth (other than dothistromin) binding to the Mab and producing a false background reading or possible contamination in the ELISA procedure. Alternatively, the dothistromin in the early samples could have been produced by mycelium fragments present in the spore suspension which were at a later, dothistromin-producing, growth stage.

Alternatively, the pattern observed could be correct and dothistromin is produced in a biphasic pattern. Repetition of the experiment and assays is required.

Overall, a rapid increase in dothistromin production (ng dothistromin per ml of broth) is observed in the late exponential phase of *D. pini* growth (Figure 4.1a), a trend which is characteristic of secondary metabolite production.

4.6 EXPRESSION ANALYSIS.

RT-PCR was performed to investigate the expression pattern of the putative dothistromin genes. Results (Figure 4.2) showed that the *D. pini* genes *dkr1*, *dte1*, *dox1* and *dtp1* are

expressed. The expression pattern of the five λCGV1 genes indicated that they were not expressed in a coregulated manner, with the *dte1* and *dox1* genes displaying a constitutive expression pattern. The expression pattern of *dkr1* is interesting, as a bright band was observed at day 8.5 (204 hrs), one sample before the large dothistromin peak. A faint band is observed for the other time points.

The faint bands observed in RT-PCR analysis for some of the genes, and lack of bands in northern hybridisation (other than the positive control) indicates that the expression level of the genes is low and RT-PCR conditions were not optimal for each reaction. Optimising RT-PCR conditions for each primer pair should be attempted.

The *ddh1* gene does not appear to be expressed, as no product was observed with the *ddh1* primers. The positive control performed with each RT-PCR reaction indicated that there was no significant RNA degradation. However, the position of the DS150ep1 primer used for *ddh1* cDNA synthesis and amplification is positioned closer to the poly (A) end of the cDNA (~150 bp) than any of the primers used in this analysis, so RNA degradation may have possibly deleted this primer site. The *ddh1* gene was found to contain a premature stop codon (Figure 3.12); any mutations in the promoter region may prevent transcription.

The constitutive expression pattern observed for *dox1* and *dte1* may indicate that these genes encode products which are also involved in primary metabolism. Alternatively, this result could suggest that aspects of dothistromin biosynthesis are controlled at the translational level (if the genes are involved in dothistromin production). Studies which altered the temporal expression of AF biosynthetic genes to a constitutive pattern had no effect on the AF production profile, indicating another level of control (Flaherty and Payne, 1997; Section 1.6.3).

To draw any solid inferences from these results, the exact relationship between these genes will need to be determined. That is, we would only expect to see coordinate expression if these genes are part of the same biosynthetic pathway. One way to determine this without gene disruption is to establish conditions which do not support dothistromin biosynthesis. Expression of genes involved in dothistromin biosynthesis

should not be able to be detected from *D. pini* grown on these conditions (if genes are transcriptionally regulated), whereas genes that are not specifically involved in dothistromin production will be expressed in both conditions.

Overall, the results presented here show that the *D. pini* genes *dkr1*, *dte1*, *dox1* and *dtp1* are expressed. These results also provide a foundation on which future expression experiments of these genes can be based. A comparison of the regulatory mechanisms between dothistromin and ST/AF biosynthesis is an obvious focal point for further research. Whether the dothistromin biosynthetic cluster contains a transcription factor similar to AFLR remains to be shown.

5. EVOLUTIONARY ASPECTS OF THE PUTATIVE DOTHISTROMIN CLUSTER.

5.1 INTRODUCTION.

Sequencing of a 13.3 kb *D. pini* genomic region identified a likely dothistromin biosynthetic cluster which exhibits similarities to ST/AF gene clusters (Chapter 3). Phylogenetic analysis was performed to investigate whether horizontal transfer of the *dkr1* region had occurred between the fungi. The GC content and codon usage patterns for the five genes were also determined (separately and combined). Overall, this analysis did not identify any direct evolutionary interrelationship between the two pathways.

5.2 SEQUENCING AND PHYLOGENETIC ANALYSIS OF THE *D. PINI* β-TUBULIN GENE, *TUB1*.

The *D. pini* β-tubulin gene, *tub1*, was isolated and partially sequenced by Bidlake (1996). In this study, fragments adjacent to the sequenced region were sub-cloned (Table 2.1; Section 2.11), sequenced along both strands (Section 2.14.1) and the sequence was organised into a single contig (Appendix 1.2 and 1.3). The *D. pini tub1* gene is an excellent positive control, as it is constitutively and highly expressed, and has been used in this study for PCR, RT-PCR and northern analysis (Chapters 3 and 4).

To investigate *D. pini* phylogeny, phylogenetic analysis was attempted using the *D. pini tub1* sequence aligned with other fungal β-tubulin coding sequences. Unfortunately, the data produced weakly supported and conflicting phylogenetic trees, indicating that it is not suited for phylogenetic purposes. The sequence at the first two nucleotide codon positions is predominantly conserved between all the sequences (therefore phylogeny could not be resolved at this level) while mutations at the third nucleotide codon position appeared to be site saturated (data was therefore uninformative and could lead to false trees; data not shown).

5.3 DETERMINATION AND COMPARISON OF GC CONTENT.

The GC content for each of the five putative dothistromin genes, the promoter and terminator regions, and the whole of λCGV1 was determined using the GCG COMPOSITION program (results illustrated in Figure 3.3). The GC content for the λCGV1 sequence is 54%. The GC content of the β-tubulin sequence was also determined (55.3%). The β-tubulin value is assumed to be representative of the *D. pini* genome and is within the expected range for fungi (34-63%; Storck and Alexopoulos, 1970). The GC content for the *A. nidulans* ST cluster is 53% (genomic GC content of 53.7%[1]), and the *A. parasiticus ver1* region is 55% GC rich (54.4%[1]). Overall, there is no significant difference between the putative dothistromin cluster and the β-tubulin GC content, and values are similar to those for *A. nidulans* and *A. parasiticus*.

5.4 DETERMINATION AND COMPARISON OF CODON USAGE.

The codon usage for the five putative dothistromin genes (separately and combined) were measured using the GCG CODONFREQUENCY program (Appendix 2.0). Results indicated a definite preference for C in the third nucleotide codon position, with 57-59 of the 61 sense codons being utilised for each of the genes, except for *dkr1*. The *dkr1* gene also has a strong preference for C in the third nucleotide codon position but utilises 51 of the 61 sense codons. The β-tubulin codon usage was also found to exhibit a strong preference for C in the third nucleotide codon position. The β-tubulin coding sequence utilised 54 of the 61 sense codons but usage was generally more biased than for the dothistromin genes. The *A. parasiticus ver1* gene showed a preference for G or C in the third codon position and utilised 56 of the 61 sense codons.

---

[1] Figure is the GC content obtained from the Codon Usage Database (codon usage and GC content tabulated from all genes for that organism submitted to Genbank; Nakamura *et al.*, 1998).

5.5 *D. PINI* PHYLOGENETIC ANALYSIS: TEST FOR HORIZONTAL TRANSFER.

5.5.1 Construction of a Species Tree.

5.5.1.1 Amplification of ITS and 5.8S rDNA Regions.

The species tree was constructed using 5.8S rDNA sequence. The *D. pini* 5.8S rDNA region was obtained by PCR amplification using the primers ITS4 and ITS5 (data not shown). The product was purified (Section 2.6.4), then directly sequenced using the same primers (Section 2.14.2). Sequence is shown in Figure 5.1.

5.5.1.2 5.8S Phylogenetic Analysis.

The *D. pini* and other fungal 5.8S rDNA sequences were aligned and analysed using PAUP* (Figure 5.2a; Section 2.19). Branch and bound parsimony was performed which created 13 different possible trees from which the strict consensus was taken (Figure 5.2aI). Bootstrap analysis showed medium-low support for this tree. These results indicate that the aligned 5.8S rDNA sequences contain a number of conflicting phylogenetic signals which do not clearly support a unique tree.

Split decomposition (Bandelt and Dress, 1992) is designed for analysing such data. For ideal data, split decomposition gives rise to a tree, whereas highly conflicting data gives rise to box-like structures (Buneman graphs). Splits Tree (Huson, 1998) was used for split decomposition analysis of the 5.8S rDNA data and produced a tree-like graph with a small amount of boxing (Figure 5.2aII). The boxing is representative of the contradictory signals. Bootstrap analysis indicated good support for the major branches of this tree. The same tree was observed when sites with a deletion or insertion were removed, further suggesting that the tree is representative of the data (data not shown).

5.5.2 Construction of a Gene Tree.

The *D. pini dkr1* nucleotide sequence (without introns) was aligned with that of the *A. parasiticus ver1* gene, *A. nidulans stcU* gene, *C. heterostrophus brn1* gene, *M. grisea thnR* gene and the *C. lagenarium thr1* gene (Section 2.19; see Table 3.1 for references).

```
gttccggaga acctgcggag ggatcattaC TGAGTGAGGG CGAAAGCCCG ACCTCCAACC 60
CTTTGTGAAC CAACTCTGTT GCTTCGGGGG CGACCCCGCC GTTTCGGCGA CGGCGCCCCC 120
GGAGGTCATC AAACACTGCA TCTTTGCGTC GGAGTCTTAA AGTAAATTTA AACAAACTT 180
TCAACAACGG ATCTCTTGGT TCTGGCATCG ATGAAGAACG CAGCGAAATG CGATAAGTAA 240
TGTGAATTGC AGAATTCAGT GAATCATCGA ATCTTTGAAC GCACATTGCG CCCCGTGGTA 300
TTCCGCGGGG CATGCCTGTT CGAGCGTCAT TTCACCACTC AAGCCTGGCT TGGTATTGGG 360
CGTCGCGGTT CCGCGCGCCT TAAAGTCTCC GGCTGAGCAG TTCGTCTCTA AGCGTTGTGG 420
CATATATTTC GCTGAAGAGT TCGGACGGCT TTTGGCCGTT AAATCTTTTA caaggttgac 480
ctcggatcag gtagggatac ccgctg                                      506
```

Figure 5.1 *Dothistroma pini* rDNA sequence.

Sequence of the product amplified using primers ITS4 and ITS5 is shown. Sequence 1-29 is partial 18S ribosomal RNA gene sequence (lower case sequence). Sequence coloured magenta (30-174) is the internal transcribed spacer 1 (ITS1). Cyan coloured sequence (175-332) indicates 5.8S ribosomal RNA. Green sequence represents ITS2 sequence (333-470). The remaining sequence (471-506) is the start of the 28S ribosomal RNA sequence.

Branch and bound parsimony was performed which produced a single, strongly supported, tree (Figure 5.2bI). To enable direct comparison to the species tree, split decomposition analysis was also performed using this data. This produced a highly supported tree-like graph, indicating the very good phylogenetic quality of this data (Figure 5.2bII). As for the species tree, removal of deletion and insertion sites did not alter the shape or support of the tree. Identical trees were also produced with analysis using the first two nucleotide codon positions or just the third nucleotide codon position (data not shown). These results confirm that the constructed tree is representative of the data.

5.5.3 Comparison of the Gene and Species Trees.

Examination and comparison of the two split decomposition trees (Figure 5.2aII and bII) clearly indicated that each is representative of the other, with a very similar branch pattern observed between the two trees. Removal from the species tree of genera not present on the gene tree confirms this (data not shown). Therefore, no clear evidence for horizontal transfer of these sequences between the fungi is contained within these trees. Horizontal transfer would have been indicated, had the *D. pini* external edge branched off the *Aspergillus* edge (or vice versa).

Figure 5.2 ***D. pini*** **Phylogenetic analysis.**

a) ***D. pini*** **Species Tree.** Phylogenetic tree using; **I.** branch and bound parsimony (PAUP*) and **II.** Split decomposition (Splits Tree; Huson, 1998). Numbers indicate bootstrap values for the respective branch. Abbreviations and sequence accession numbers for the 5.8S sequences used in the analysis are shown below. See Appendix 3.0 for the taxonomic classification of the fungi below.

Name	**Abbreviation**	**Accession number**
Dothistroma pini	Dpini	This study
Mycosphaerella graminicola	Mgram	U77363
Aspergillus parasiticus	Apara	AF027862
Aspergillus flavus	Aflav	AF027863
Aspergillus nidulans	Anid	L76746
Aspergillus niger	Aniger	U65306
Alternaria alternata.	Aalt	X17454
Magnaporthe grisea	Mgris	U17329
Colletotrichum graminicola	Cgram	AF059676
Pyrenophora teres	Pteres	Y08744
Filobasidium uniguttulatum	Funig	M94520
Neurospora crassa	Ncras	X02447
Filobasidiella neoformans	Fneo	M94518
Cystofilobasidium capitatum	Ccapit	M94512

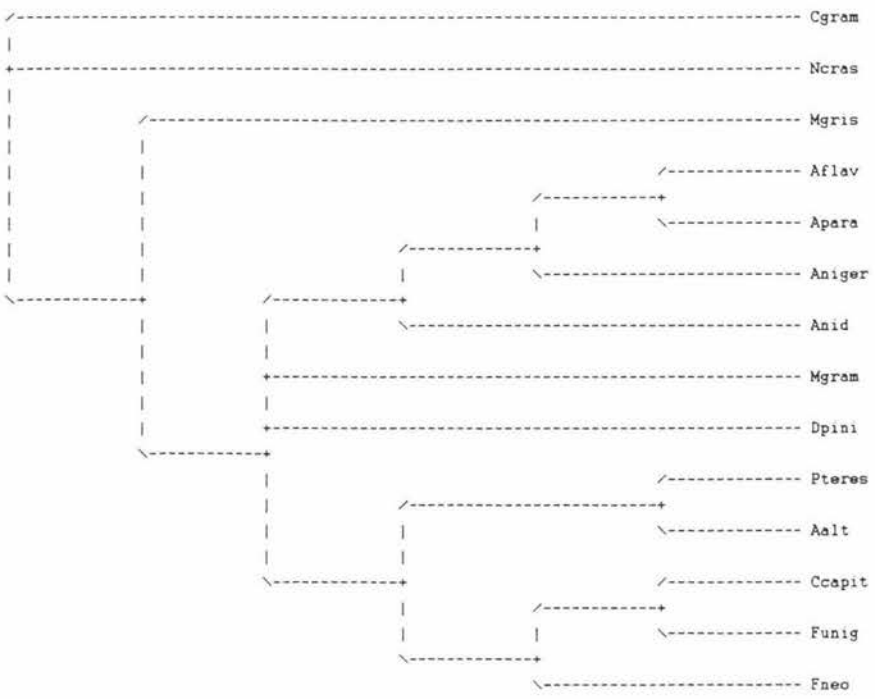
b) ***D. pini dkr1*** **Gene Tree.** Phylogenetic tree using; **I.** branch and bound parsimony (PAUP*) and **II.** Split decomposition (Splits Tree; Huson, 1998). Abbreviations and sequence reference for the 5.8S sequences used in the analysis are shown below. Numbers represent the bootstrap value for the respective branch. See Table 3.1 for references. Abbreviations are:

Dpiniver1	*D. pini dkr1.*
AnstcU	*A. nidulans stcU.*
Apver1	*A. parasiticus ver1.*
ChBrn1	*C. heterostrophus*
MgthnR	*M. grisea thnR*
ClTHR1	*C. lagenarium thr1*

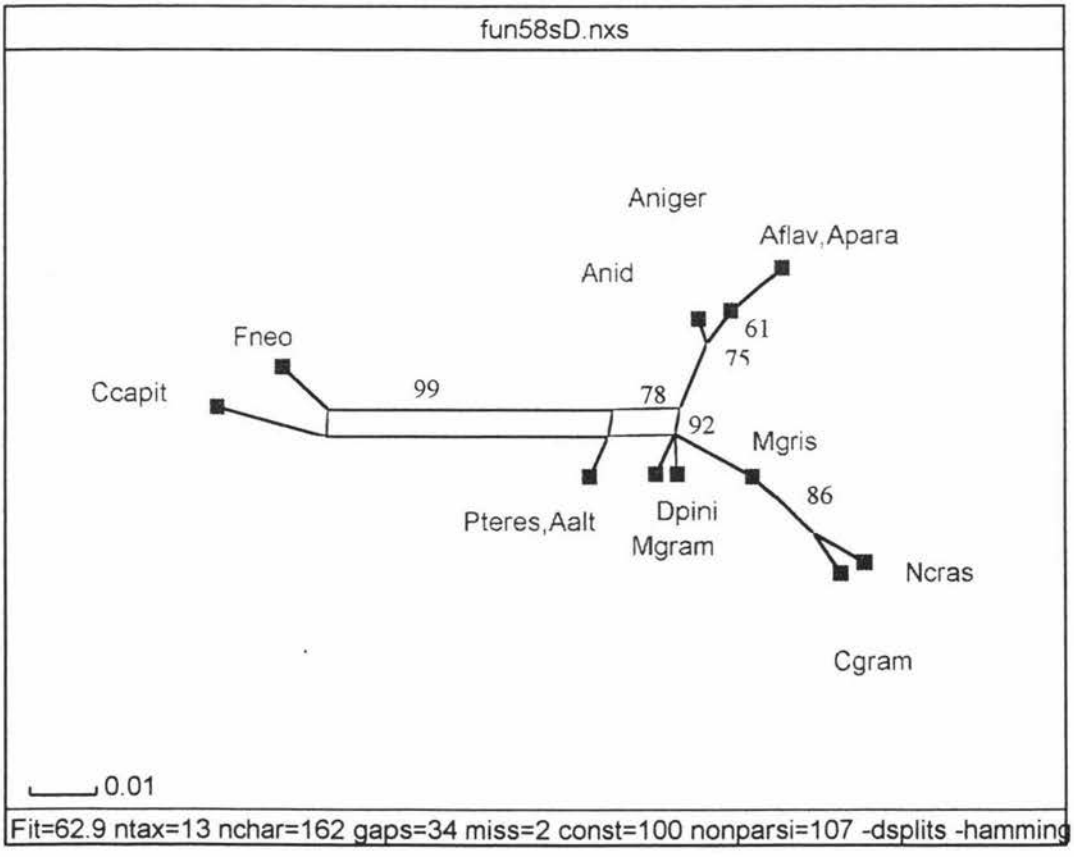
a)

I.

Branch-and-bound Parsimony tree
Strict consensus of 13 trees:



II.

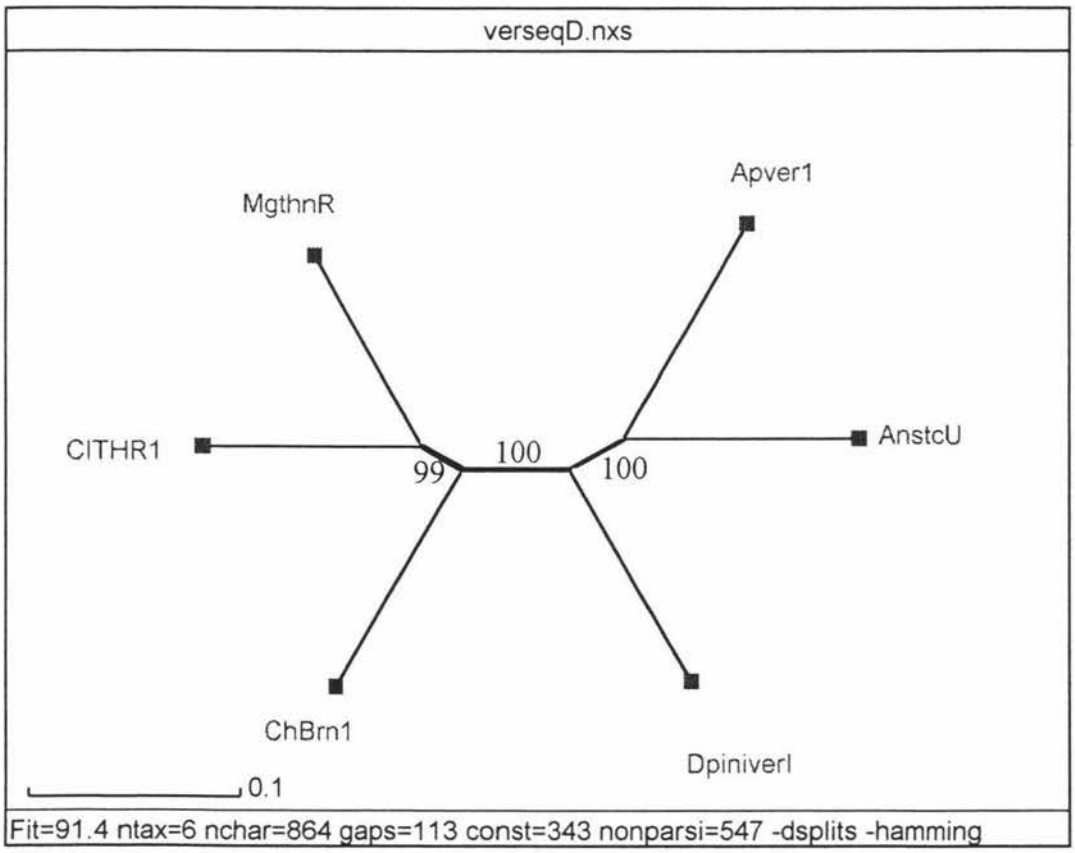


b)

I

```
Bootstrap method with branch-and-bound Parsimony

                                                    /------------------------- MgthnR(2)
/-----------------------100-------------------------+
|                                                   \------------------------- ClTHR1(6)
|
|                                                   /------------------------- AnstcU(1)
|                         /-----------100-----------+
|                         |                         \------------------------- Apver1(3)
+----------100-----------+
|                         \------------------------------------------------- DpiniverI(5)
|
\------------------------------------------------------------------------------ ChBrn1(4)
```

II



CHAPTER FIVE. DISCUSSION.

5.6 THE β-TUBULIN GENE.

The *D. pini* β-tubulin gene, *tub1*, was sequenced and was used throughout this study as a positive control. The *tub1* sequence is located outside the putative dothistromin cluster and provides a sequences which is considered to be representative of the *D. pini* genome. Unfortunately the β-tubulin coding sequence was unsuited for phylogenetic analysis. The *tub1* gene is constitively expressed and the *tub1* promoter will be a useful tool for future investigations which require constitively expressed constructs.

5.7 EVOLUTION OF THE PUTATIVE DOTHISTROMIN CLUSTER.

The *D. pini dkr1* gene exhibits high nucleotide and deduced amino acid sequence identity to ketoreductases involved in ST/AF and melanin production. Genes located within close proximity to *dkr1* (*dox1* and *dte1*) also show significant identity to genes involved in ST/AF biosynthesis. These results suggest that these sequences may have a close evolutionary relationship (Section 3.11.1.1).

Dothistroma and *Aspergillus* are classified in the same class (Ascomycotina) but are in different sub-classes (loculoascomycetes and plectomycetes respectively; Appendix 3), indicating a somewhat distant relationship. Considering that dothistromin and aflatoxin genes are not required for growth of the fungus, and that the toxins are seemingly involved in different roles (Section 3.11.1.1), it is interesting that the *dkr1*, *ver1* and *stcU* genes display such high identity.

To investigate the hypothesis that horizontal transmission of the *ver1/dkr1* sequence (and possibly the whole cluster) has occurred between the fungi, phylogenetic analysis was performed. Hypotheses of horizontal transmission can be established if it can be demonstrated that there is strongly supported conflict between the phylogeny of the element that is thought to have been transmitted horizontally and the "species" phylogeny of the organisms which bear the element (Section 1.7.2; Li, 1997 and references therein). This type of analysis has been used successfully in other examples for filamentous fungi (Hibbett, 1996; Buades and Moya, 1996).

The phylogeny of the element under investigation, the *dkr1* gene, was determined. A strongly supported "gene" tree was constructed using the nucleotide coding sequences of *dkr1*, *ver1*, *stcU*, *brn1*, *thnR*, and *thr1* (Figure 5.2b). A species tree was constructed using the *D. pini* and other fungal 5.8S rDNA sequences (Figure 5.2a). The *D. pini* 5.8S (also ITS1 and ITS2) sequence was obtained by direct sequencing of the respective region amplified using the ITS4 and ITS5 primers (Figure 5.1).

Comparison of the gene and species trees identified no obvious conflict between the two, indicating that horizontal transmission of the *dkr1/ver1* region (and possibly the cluster) has not occurred. However, these results should be viewed with caution. The 5.8S rDNA sequences were shown to contain conflicting signals, with split decomposition analysis being performed so to alleviate these problems. Split decomposition is not the best method for assessing incongruence between two data sets (Hibbett, 1996), therefore it is not the optimal vector in which to evaluate potential cases of horizontal transmission.

Based on phylogenetic analysis similar to above, Buades and Moya (1996) proposed the horizontal transmission of the IPN-synthetase gene from bacteria to fungi. In their study, comparison of gene and species trees (constructed using maximum likelihood or parsimony methods) revealed some conflicts between the two. However, the major difference between the trees were the branch lengths. The branch lengths for the gene tree were significantly shorter than those for the species tree, indicating a eariler divergence time than that represented in the species tree (therefore suggesting a horizontal transfer event). Whether differences in branch lengths are present in the *dkr1* gene tree can not be resolved with the 5.8S data (and Split decomposition analysis). The 18S rDNA sequence may provide better data for the construction of the species tree, as it is larger than the 5.8S rDNA region and therefore should contain more informative sites. Analysis with 18S rDNA sequences should be attempted.

Certain *Cercospora* species have been identified which are able to produce dothistromin (Assante et al., 1977). Likewise, many *Aspergillus* species have been shown to produce intermediates of the AF pathway, suggesting that these fungi contain functional sections of the pathway and that the progenitor *Aspergillus* strain also contained the ST/AF pathway (or a version of) (Trail *et al.*, 1995). Physical clustering may suggest that the

progenitor strain obtained the pathway intact via horizontal transfer from some other organism. Alternatively, the pathway may have evolved from a pre-existing fungal polyketide pathway (possibly from the melanin biosynthetic gene cluster). This could account for the high similarity of the dothistromin/AF ketoreductases and the melanin reductases.

Many asexual fungal pathogens demonstrate variable karyotypes, and studies with Australian biotypes of *Colletotrichum gloeosporioides* has detected horizontal transfer of entire nuclear chromosomes between different biotypes (Masel, *et al.*, 1996). Transmission of a linear autonomously replicating transformation vector between two *C. gloeosporioides* biotypes has also been demonstrated in mycelial paring experiments (Masel, *et al.*, 1996). Horizontal transfer of a mitochondrial plasmid has also been shown between genetically incompatible fungal species (Kempken, 1995). These examples illustrate that horizontal gene transfer between fungi is possible, although the mechanisms of which are unknown.

SUMMARY AND FUTURE DIRECTIONS.

In our laboratory, the molecular genetics of dothistromin biosynthesis is being investigated. In this study, five genes were identified (*dkr1*, *dox1*, *dte1*, *dtp1*, and *ddh1*), over a 13.3 kb *D. pini* genomic DNA region, which may comprise a section of a dothistromin biosynthetic gene cluster (Table 3.1, Figure 3.3). The deduced amino acid sequence of the *D. pini dkr1* gene displays high sequence identity to ketoreductases involved in ST/AF and melanin biosynthesis. The *dox1* gene exhibits similarities to oxidases, including STCC which is involved in ST production. The *dte1* gene encodes a product with significant sequence similarity to the thioesterase domain of PKSs involved in ST/AF production and conidial pigmentation, and is intriguing as it may comprise part of a type II PKS. The deduced amino acid sequence of the *dtp1* gene contains a high level of hydrophobic residues and shows strong similarities to transmembrane transport proteins. Because of this similarity and the location of the *dtp1* gene, DTP1 is proposed to be a dothistromin efflux pump. The *ddh1* gene shows similarities to bacterial polysaccharide dehydrogenases, however the *ddh1* coding sequence contains a premature stop codon. Southern analysis confirmed genomic clustering of the genes and indicated that each gene is present in a single copy. Expression analysis indicated that each of the genes, except *ddh1*, is expressed.

Results from this study, and comparison with the ST/AF systems, indicates that we have identified a portion of a likely dothistromin biosynthetic cluster. Effort can now be focused onto the characterisation of the genes identified and extension of this sequence. Characterisation should include targeted gene disruption of each gene (especially *dkr1*). This will elucidate the function of the gene product and confirm whether it is involved in dothistromin biosynthesis. Dothistromin-minus mutants can then be used to evaluate the role of dothistromin in pathogenesis. Extension of the λCGV1 sequence will identify further genes which may be part of the cluster. The regions surrounding the λCGV1 sequence can be isolated by chromosome walking (in both directions initially) using end λCGV1 fragments as probes. Alternatively, fragments not completely contained on the λCGV1 ends (size determined by Southern analysis) could be cloned and sequenced. For example, Southern analysis detected a 17.6 kb *Sal*I fragment flanking the right lambda arm, of which approximately 11 kb is not contained within λCGV1. This region could be

obtained by digesting genomic DNA (with *Sal*I), isolating the DNA within the expect size range (16-18 kb), cloning these fragments using a pUC vector (depending on size) and screening *E. coli* transformants using a probe from λCGV1 covering a section of the fragment (eg the *Eco*RI 2.1 kb fragment).

The relationship between the dothistromin and ST/AF biosynthetic pathways is very interesting and should be investigated further. Determining whether the *D. pini dkr1* gene is able to complement an *A. parasiticus ver1* mutant would give a very good indication of the relationship between these pathways. The association should also be investigated from an evolution view point, and phylogenetic analysis initiated here should be continued. This analysis could be extended to incorporate the construction of gene trees for other sequences (particularly *dte1* and *ddh1*). Critical to this analysis is the construction of a clear species tree. This could be achieved using 18S rDNA data or possibly mitochondrial rDNA sequence, both of which can be easily isolated by PCR (White, *et al.*, 1990).

This study touched on the expression pattern of the putative dothistromin genes. This analysis provides a good starting point for further, more detailed investigation. Conditions which support the growth of the fungus but not dothistromin production should be elucidated. These conditions will facilitate the identification of genes specifically expressed in dothistromin producing conditions. Central to analysing the regulatory mechanisms controlling dothistromin biosynthesis will be the isolation of an *aflR*-like gene.

The fungus *Cercospora smilacis* also produces dothistromin (Assante *et al.*, 1977). Southern blot analysis containing *C. smilacis* DNA, hybridised with putative dothistromin genes, should be performed to investigate the level of similarity between the two pathways and the genomic organisation (that is, whether the pathway genes in *C. smilacis* are also clustered). If possible, sequences of *C. smilacis* dothistromin genes should be incorporated into the phylogenetic analysis.

Appendices

Appendix 1.0 Sequence Data.

Because of the volume of sequence, the following are presented as files stored on the enclosed disk.

Appendix 1.1 GCG PRETTYOUT From λCGV1 Contig.

Appendix 1.2 GCG PRETTYOUT From β-tubulin Contig.

Appendix 1.3 GCG BIGPICTURE From β-tubulin Contig.

Appendix 2.0 Codon Usage Tables.

Combined Codon Usage Table for *ddh1*, *dkr1*, *dox1*, *dtp1* and *dte1*.

AmAcid	Codon	Number	/1000	Fraction
Gly	GGG	15.00	7.44	0.09
Gly	GGA	33.00	16.36	0.19
Gly	GGT	47.00	23.30	0.27
Gly	GGC	78.00	38.67	0.45
Glu	GAG	47.00	23.30	0.65
Glu	GAA	25.00	12.39	0.35
Asp	GAT	39.00	19.34	0.42
Asp	GAC	54.00	26.77	0.58
Val	GTG	29.00	14.38	0.22
Val	GTA	13.00	6.45	0.10
Val	GTT	16.00	7.93	0.12
Val	GTC	76.00	37.68	0.57
Ala	GCG	39.00	19.34	0.18
Ala	GCA	55.00	27.27	0.25
Ala	GCT	44.00	21.81	0.20
Ala	GCC	84.00	41.65	0.38
Arg	AGG	5.00	2.48	0.06
Arg	AGA	18.00	8.92	0.20
Ser	AGT	6.00	2.97	0.04
Ser	AGC	24.00	11.90	0.17
Lys	AAG	55.00	27.27	0.69
Lys	AAA	25.00	12.39	0.31
Asn	AAT	15.00	7.44	0.20
Asn	AAC	61.00	30.24	0.80
Met	ATG	48.00	23.80	1.00
Ile	ATA	6.00	2.97	0.04
Ile	ATT	29.00	14.38	0.21
Ile	ATC	100.00	49.58	0.74
Thr	ACG	20.00	9.92	0.16
Thr	ACA	24.00	11.90	0.20
Thr	ACT	26.00	12.89	0.21
Thr	ACC	52.00	25.78	0.43
Trp	TGG	24.00	11.90	1.00
End	TGA	3.00	1.49	0.50
Cys	TGT	9.00	4.46	0.30
Cys	TGC	21.00	10.41	0.70
End	TAG	2.00	0.99	0.33
End	TAA	1.00	0.50	0.17
Tyr	TAT	16.00	7.93	0.24
Tyr	TAC	51.00	25.29	0.76
Leu	TTG	35.00	17.35	0.20
Leu	TTA	3.00	1.49	0.02
Phe	TTT	25.00	12.39	0.25
Phe	TTC	75.00	37.18	0.75
Ser	TCG	28.00	13.88	0.19
Ser	TCA	23.00	11.40	0.16
Ser	TCT	23.00	11.40	0.16

Ser	TCC	40.00	19.83	0.28
Arg	CGG	10.00	4.96	0.11
Arg	CGA	10.00	4.96	0.11
Arg	CGT	14.00	6.94	0.16
Arg	CGC	33.00	16.36	0.37
Gln	CAG	41.00	20.33	0.53
Gln	CAA	36.00	17.85	0.47
His	CAT	11.00	5.45	0.31
His	CAC	25.00	12.39	0.69
Leu	CTG	41.00	20.33	0.23
Leu	CTA	7.00	3.47	0.04
Leu	CTT	26.00	12.89	0.15
Leu	CTC	64.00	31.73	0.36
Pro	CCG	31.00	15.37	0.28
Pro	CCA	27.00	13.39	0.24
Pro	CCT	18.00	8.92	0.16
Pro	CCC	36.00	17.85	0.32

Codon Usage Table for *ddh1*.

AmAcid	Codon	Number	/1000	Fraction
Gly	GGG	5.00	10.64	0.20
Gly	GGA	7.00	14.89	0.28
Gly	GGT	1.00	2.13	0.04
Gly	GGC	12.00	25.53	0.48
Glu	GAG	11.00	23.40	0.55
Glu	GAA	9.00	19.15	0.45
Asp	GAT	10.00	21.28	0.40
Asp	GAC	15.00	31.91	0.60
Val	GTG	11.00	23.40	0.30
Val	GTA	3.00	6.38	0.08
Val	GTT	4.00	8.51	0.11
Val	GTC	19.00	40.43	0.51
Ala	GCG	9.00	19.15	0.23
Ala	GCA	7.00	14.89	0.18
Ala	GCT	8.00	17.02	0.21
Ala	GCC	15.00	31.91	0.38
Arg	AGG	0.00	0.00	0.00
Arg	AGA	8.00	17.02	0.28
Ser	AGT	0.00	0.00	0.00
Ser	AGC	8.00	17.02	0.19
Lys	AAG	8.00	17.02	0.67
Lys	AAA	4.00	8.51	0.33
Asn	AAT	3.00	6.38	0.27
Asn	AAC	8.00	17.02	0.73
Met	ATG	13.00	27.66	1.00
Ile	ATA	3.00	6.38	0.10
Ile	ATT	6.00	12.77	0.20
Ile	ATC	21.00	44.68	0.70
Thr	ACG	7.00	14.89	0.23
Thr	ACA	8.00	17.02	0.26
Thr	ACT	7.00	14.89	0.23
Thr	ACC	9.00	19.15	0.29
Trp	TGG	4.00	8.51	1.00
End	TGA	2.00	4.26	1.00
Cys	TGT	0.00	0.00	0.00
Cys	TGC	7.00	14.89	1.00
End	TAG	0.00	0.00	0.00
End	TAA	0.00	0.00	0.00
Tyr	TAT	8.00	17.02	0.47
Tyr	TAC	9.00	19.15	0.53
Leu	TTG	5.00	10.64	0.14
Leu	TTA	1.00	2.13	0.03
Phe	TTT	4.00	8.51	0.29
Phe	TTC	10.00	21.28	0.71
Ser	TCG	7.00	14.89	0.16
Ser	TCA	8.00	17.02	0.19
Ser	TCT	11.00	23.40	0.26
Ser	TCC	9.00	19.15	0.21

Arg	CGG	5.00	10.64	0.17
Arg	CGA	5.00	10.64	0.17
Arg	CGT	4.00	8.51	0.14
Arg	CGC	7.00	14.89	0.24
Gln	CAG	12.00	25.53	0.50
Gln	CAA	12.00	25.53	0.50
His	CAT	3.00	6.38	0.23
His	CAC	10.00	21.28	0.77
Leu	CTG	7.00	14.89	0.20
Leu	CTA	2.00	4.26	0.06
Leu	CTT	5.00	10.64	0.14
Leu	CTC	15.00	31.91	0.43
Pro	CCG	11.00	23.40	0.28
Pro	CCA	3.00	6.38	0.08
Pro	CCT	7.00	14.89	0.18
Pro	CCC	18.00	38.30	0.46

Codon Usage Table for *dkr1*.

AmAcid	Codon	Number	/1000	Fraction
Gly	GGG	2.00	7.58	0.07
Gly	GGA	2.00	7.58	0.07
Gly	GGT	14.00	53.03	0.48
Gly	GGC	11.00	41.67	0.38
Glu	GAG	11.00	41.67	0.92
Glu	GAA	1.00	3.79	0.08
Asp	GAT	6.00	22.73	0.32
Asp	GAC	13.00	49.24	0.68
Val	GTG	1.00	3.79	0.03
Val	GTA	1.00	3.79	0.03
Val	GTT	4.00	15.15	0.14
Val	GTC	23.00	87.12	0.79
Ala	GCG	6.00	22.73	0.17
Ala	GCA	4.00	15.15	0.11
Ala	GCT	8.00	30.30	0.23
Ala	GCC	17.00	64.39	0.49
Arg	AGG	0.00	0.00	0.00
Arg	AGA	0.00	0.00	0.00
Ser	AGT	1.00	3.79	0.07
Ser	AGC	1.00	3.79	0.07
Lys	AAG	13.00	49.24	0.87
Lys	AAA	2.00	7.58	0.13
Asn	AAT	0.00	0.00	0.00
Asn	AAC	9.00	34.09	1.00
Met	ATG	5.00	18.94	1.00
Ile	ATA	0.00	0.00	0.00
Ile	ATT	6.00	22.73	0.29
Ile	ATC	15.00	56.82	0.71
Thr	ACG	0.00	0.00	0.00
Thr	ACA	1.00	3.79	0.09
Thr	ACT	4.00	15.15	0.36
Thr	ACC	6.00	22.73	0.55
Trp	TGG	2.00	7.58	1.00
End	TGA	0.00	0.00	0.00
Cys	TGT	1.00	3.79	0.33
Cys	TGC	2.00	7.58	0.67
End	TAG	1.00	3.79	1.00
End	TAA	0.00	0.00	0.00
Tyr	TAT	0.00	0.00	0.00
Tyr	TAC	5.00	18.94	1.00
Leu	TTG	1.00	3.79	0.08
Leu	TTA	0.00	0.00	0.00
Phe	TTT	4.00	15.15	0.36
Phe	TTC	7.00	26.52	0.64
Ser	TCG	4.00	15.15	0.27
Ser	TCA	1.00	3.79	0.07
Ser	TCT	4.00	15.15	0.27
Ser	TCC	4.00	15.15	0.27
Arg	CGG	0.00	0.00	0.00
Arg	CGA	3.00	11.36	0.20

Arg	CGT	5.00	18.94	0.33
Arg	CGC	7.00	26.52	0.47
Gln	CAG	2.00	7.58	0.40
Gln	CAA	3.00	11.36	0.60
His	CAT	1.00	3.79	0.25
His	CAC	3.00	11.36	0.75
Leu	CTG	2.00	7.58	0.15
Leu	CTA	1.00	3.79	0.08
Leu	CTT	3.00	11.36	0.23
Leu	CTC	6.00	22.73	0.46
Pro	CCG	2.00	7.58	0.40
Pro	CCA	0.00	0.00	0.00
Pro	CCT	0.00	0.00	0.00
Pro	CCC	3.00	11.36	0.60

Codon Usage Table for *dox1*.

AmAcid	Codon	Number	/1000	Fraction	..
Gly	GGG	5.00	12.05	0.11	
Gly	GGA	9.00	21.69	0.20	
Gly	GGT	10.00	24.10	0.22	
Gly	GGC	22.00	53.01	0.48	
Glu	GAG	3.00	7.23	0.33	
Glu	GAA	6.00	14.46	0.67	
Asp	GAT	12.00	28.92	0.57	
Asp	GAC	9.00	21.69	0.43	
Val	GTG	9.00	21.69	0.38	
Val	GTA	4.00	9.64	0.17	
Val	GTT	3.00	7.23	0.12	
Val	GTC	8.00	19.28	0.33	
Ala	GCG	8.00	19.28	0.20	
Ala	GCA	14.00	33.73	0.35	
Ala	GCT	5.00	12.05	0.12	
Ala	GCC	13.00	31.33	0.32	
Arg	AGG	2.00	4.82	0.25	
Arg	AGA	0.00	0.00	0.00	
Ser	AGT	0.00	0.00	0.00	
Ser	AGC	8.00	19.28	0.24	
Lys	AAG	11.00	26.51	0.69	
Lys	AAA	5.00	12.05	0.31	
Asn	AAT	7.00	16.87	0.23	
Asn	AAC	23.00	55.42	0.77	
Met	ATG	9.00	21.69	1.00	
Ile	ATA	0.00	0.00	0.00	
Ile	ATT	5.00	12.05	0.28	
Ile	ATC	13.00	31.33	0.72	
Thr	ACG	7.00	16.87	0.24	
Thr	ACA	2.00	4.82	0.07	
Thr	ACT	6.00	14.46	0.21	
Thr	ACC	14.00	33.73	0.48	
Trp	TGG	4.00	9.64	1.00	
End	TGA	0.00	0.00	0.00	
Cys	TGT	1.00	2.41	0.17	
Cys	TGC	5.00	12.05	0.83	
End	TAG	1.00	2.41	1.00	
End	TAA	0.00	0.00	0.00	
Tyr	TAT	2.00	4.82	0.11	
Tyr	TAC	16.00	38.55	0.89	
Leu	TTG	12.00	28.92	0.36	
Leu	TTA	0.00	0.00	0.00	
Phe	TTT	2.00	4.82	0.08	
Phe	TTC	24.00	57.83	0.92	
Ser	TCG	9.00	21.69	0.26	
Ser	TCA	4.00	9.64	0.12	
Ser	TCT	4.00	9.64	0.12	
Ser	TCC	9.00	21.69	0.26	
Arg	CGG	0.00	0.00	0.00	
Arg	CGA	0.00	0.00	0.00	
Arg	CGT	2.00	4.82	0.25	
Arg	CGC	4.00	9.64	0.50	

Gln	CAG	8.00	19.28	0.53
Gln	CAA	7.00	16.87	0.47
His	CAT	4.00	9.64	0.57
His	CAC	3.00	7.23	0.43
Leu	CTG	6.00	14.46	0.18
Leu	CTA	1.00	2.41	0.03
Leu	CTT	6.00	14.46	0.18
Leu	CTC	8.00	19.28	0.24
Pro	CCG	6.00	14.46	0.29
Pro	CCA	8.00	19.28	0.38
Pro	CCT	5.00	12.05	0.24
Pro	CCC	2.00	4.82	0.10

Codon Usage Table for *dtp1*.

AmAcid	Codon	Number	/1000	Fraction
Gly	GGG	0.00	0.00	0.00
Gly	GGA	13.00	23.85	0.25
Gly	GGT	17.00	31.19	0.33
Gly	GGC	21.00	38.53	0.41
Glu	GAG	8.00	14.68	0.80
Glu	GAA	2.00	3.67	0.20
Asp	GAT	7.00	12.84	0.54
Asp	GAC	6.00	11.01	0.46
Val	GTG	6.00	11.01	0.19
Val	GTA	4.00	7.34	0.13
Val	GTT	5.00	9.17	0.16
Val	GTC	16.00	29.36	0.52
Ala	GCG	13.00	23.85	0.18
Ala	GCA	21.00	38.53	0.28
Ala	GCT	17.00	31.19	0.23
Ala	GCC	23.00	42.20	0.31
Arg	AGG	0.00	0.00	0.00
Arg	AGA	7.00	12.84	0.44
Ser	AGT	4.00	7.34	0.13
Ser	AGC	5.00	9.17	0.17
Lys	AAG	16.00	29.36	0.80
Lys	AAA	4.00	7.34	0.20
Asn	AAT	1.00	1.83	0.06
Asn	AAC	16.00	29.36	0.94
Met	ATG	13.00	23.85	1.00
Ile	ATA	2.00	3.67	0.04
Ile	ATT	8.00	14.68	0.17
Ile	ATC	37.00	67.89	0.79
Thr	ACG	5.00	9.17	0.12
Thr	ACA	10.00	18.35	0.25
Thr	ACT	9.00	16.51	0.22
Thr	ACC	16.00	29.36	0.40
Trp	TGG	10.00	18.35	1.00
End	TGA	0.00	0.00	0.00
Cys	TGT	3.00	5.50	0.38
Cys	TGC	5.00	9.17	0.62
End	TAG	0.00	0.00	0.00
End	TAA	1.00	1.83	1.00
Tyr	TAT	4.00	7.34	0.24
Tyr	TAC	13.00	23.85	0.76
Leu	TTG	12.00	22.02	0.18
Leu	TTA	1.00	1.83	0.01
Phe	TTT	10.00	18.35	0.27
Phe	TTC	27.00	49.54	0.73
Ser	TCG	7.00	12.84	0.23
Ser	TCA	5.00	9.17	0.17
Ser	TCT	2.00	3.67	0.07
Ser	TCC	7.00	12.84	0.23
Arg	CGG	2.00	3.67	0.12
Arg	CGA	2.00	3.67	0.12
Arg	CGT	1.00	1.83	0.06
Arg	CGC	4.00	7.34	0.25
Gln	CAG	15.00	27.52	0.75
Gln	CAA	5.00	9.17	0.25

His	CAT	1.00	1.83	0.33
His	CAC	2.00	3.67	0.67
Leu	CTG	22.00	40.37	0.33
Leu	CTA	3.00	5.50	0.04
Leu	CTT	11.00	20.18	0.16
Leu	CTC	18.00	33.03	0.27
Pro	CCG	6.00	11.01	0.30
Pro	CCA	9.00	16.51	0.45
Pro	CCT	3.00	5.50	0.15
Pro	CCC	2.00	3.67	0.10

Codon Usage Table for *dte1*.

AmAcid	Codon	Number	/1000	Fraction
Gly	GGG	3.00	9.29	0.14
Gly	GGA	2.00	6.19	0.09
Gly	GGT	5.00	15.48	0.23
Gly	GGC	12.00	37.15	0.55
Glu	GAG	14.00	43.34	0.67
Glu	GAA	7.00	21.67	0.33
Asp	GAT	4.00	12.38	0.27
Asp	GAC	11.00	34.06	0.73
Val	GTG	2.00	6.19	0.15
Val	GTA	1.00	3.10	0.08
Val	GTT	0.00	0.00	0.00
Val	GTC	10.00	30.96	0.77
Ala	GCG	3.00	9.29	0.09
Ala	GCA	9.00	27.86	0.26
Ala	GCT	6.00	18.58	0.18
Ala	GCC	16.00	49.54	0.47
Arg	AGG	3.00	9.29	0.14
Arg	AGA	3.00	9.29	0.14
Ser	AGT	1.00	3.10	0.05
Ser	AGC	2.00	6.19	0.09
Lys	AAG	7.00	21.67	0.41
Lys	AAA	10.00	30.96	0.59
Asn	AAT	4.00	12.38	0.44
Asn	AAC	5.00	15.48	0.56
Met	ATG	8.00	24.77	1.00
Ile	ATA	1.00	3.10	0.05
Ile	ATT	4.00	12.38	0.21
Ile	ATC	14.00	43.34	0.74
Thr	ACG	1.00	3.10	0.09
Thr	ACA	3.00	9.29	0.27
Thr	ACT	0.00	0.00	0.00
Thr	ACC	7.00	21.67	0.64
Trp	TGG	4.00	12.38	1.00
End	TGA	1.00	3.10	1.00
Cys	TGT	4.00	12.38	0.67
Cys	TGC	2.00	6.19	0.33
End	TAG	0.00	0.00	0.00
End	TAA	0.00	0.00	0.00
Tyr	TAT	2.00	6.19	0.20
Tyr	TAC	8.00	24.77	0.80
Leu	TTG	5.00	15.48	0.18
Leu	TTA	1.00	3.10	0.04
Phe	TTT	5.00	15.48	0.42
Phe	TTC	7.00	21.67	0.58
Ser	TCG	1.00	3.10	0.05
Ser	TCA	5.00	15.48	0.23
Ser	TCT	2.00	6.19	0.09
Ser	TCC	11.00	34.06	0.50
Arg	CGG	3.00	9.29	0.14
Arg	CGA	0.00	0.00	0.00
Arg	CGT	2.00	6.19	0.09
Arg	CGC	11.00	34.06	0.50
Gln	CAG	4.00	12.38	0.31
Gln	CAA	9.00	27.86	0.69
His	CAT	2.00	6.19	0.22
His	CAC	7.00	21.67	0.78

Leu	CTG	4.00	12.38	0.14
Leu	CTA	0.00	0.00	0.00
Leu	CTT	1.00	3.10	0.04
Leu	CTC	17.00	52.63	0.61
Pro	CCG	6.00	18.58	0.22
Pro	CCA	7.00	21.67	0.26
Pro	CCT	3.00	9.29	0.11
Pro	CCC	11.00	34.06	0.41

Codon Usage Table for *tub1*.

AmAcid	Codon	Number	/1000	Fraction
Gly	GGG	1.00	2.23	0.03
Gly	GGA	2.00	4.46	0.05
Gly	GGT	18.00	40.18	0.46
Gly	GGC	18.00	40.18	0.46
Glu	GAG	33.00	73.66	0.92
Glu	GAA	3.00	6.70	0.08
Asp	GAT	5.00	11.16	0.23
Asp	GAC	17.00	37.95	0.77
Val	GTG	5.00	11.16	0.14
Val	GTA	0.00	0.00	0.00
Val	GTT	4.00	8.93	0.11
Val	GTC	26.00	58.04	0.74
Ala	GCG	3.00	6.70	0.10
Ala	GCA	5.00	11.16	0.17
Ala	GCT	9.00	20.09	0.31
Ala	GCC	12.00	26.79	0.41
Arg	AGG	1.00	2.23	0.04
Arg	AGA	0.00	0.00	0.00
Ser	AGT	1.00	2.23	0.03
Ser	AGC	3.00	6.70	0.10
Lys	AAG	13.00	29.02	1.00
Lys	AAA	0.00	0.00	0.00
Asn	AAT	3.00	6.70	0.12
Asn	AAC	23.00	51.34	0.88
Met	ATG	19.00	42.41	1.00
Ile	ATA	1.00	2.23	0.07
Ile	ATT	4.00	8.93	0.29
Ile	ATC	9.00	20.09	0.64
Thr	ACG	4.00	8.93	0.15
Thr	ACA	0.00	0.00	0.00
Thr	ACT	9.00	20.09	0.35
Thr	ACC	13.00	29.02	0.50
Trp	TGG	4.00	8.93	1.00
End	TGA	0.00	0.00	0.00
Cys	TGT	2.00	4.46	0.25
Cys	TGC	6.00	13.39	0.75
End	TAG	1.00	2.23	1.00
End	TAA	0.00	0.00	0.00
Tyr	TAT	3.00	6.70	0.21
Tyr	TAC	11.00	24.55	0.79
Leu	TTG	5.00	11.16	0.15
Leu	TTA	1.00	2.23	0.03
Phe	TTT	1.00	2.23	0.04
Phe	TTC	24.00	53.57	0.96
Ser	TCG	6.00	13.39	0.20
Ser	TCA	1.00	2.23	0.03
Ser	TCT	2.00	4.46	0.07
Ser	TCC	17.00	37.95	0.57
Arg	CGG	0.00	0.00	0.00
Arg	CGA	1.00	2.23	0.04
Arg	CGT	9.00	20.09	0.39
Arg	CGC	12.00	26.79	0.52
Gln	CAG	19.00	42.41	0.86
Gln	CAA	3.00	6.70	0.14
His	CAT	2.00	4.46	0.22
His	CAC	7.00	15.63	0.78
Leu	CTG	5.00	11.16	0.15

Leu	CTA	0.00	0.00	0.00
Leu	CTT	5.00	11.16	0.15
Leu	CTC	18.00	40.18	0.53
Pro	CCG	0.00	0.00	0.00
Pro	CCA	16.00	35.71	0.84
Pro	CCT	2.00	4.46	0.11
Pro	CCC	1.00	2.23	0.05

Codon Usage Table for the *Aspergillus nidulans stcU* gene..

AmAcid	Codon	Number	/1000	Fraction
Gly	GGG	5.00	18.66	0.17
Gly	GGA	7.00	26.12	0.23
Gly	GGT	10.00	37.31	0.33
Gly	GGC	8.00	29.85	0.27
Glu	GAG	2.00	7.46	0.40
Glu	GAA	3.00	11.19	0.60
Asp	GAT	11.00	41.04	0.69
Asp	GAC	5.00	18.66	0.31
Val	GTG	3.00	11.19	0.12
Val	GTA	5.00	18.66	0.20
Val	GTT	4.00	14.93	0.16
Val	GTC	13.00	48.51	0.52
Ala	GCG	6.00	22.39	0.23
Ala	GCA	5.00	18.66	0.19
Ala	GCT	6.00	22.39	0.23
Ala	GCC	9.00	33.58	0.35
Arg	AGG	2.00	7.46	0.07
Arg	AGA	2.00	7.46	0.07
Ser	AGT	1.00	3.73	0.05
Ser	AGC	4.00	14.93	0.18
Lys	AAG	5.00	18.66	0.56
Lys	AAA	4.00	14.93	0.44
Asn	AAT	1.00	3.73	0.20
Asn	AAC	4.00	14.93	0.80
Met	ATG	3.00	11.19	1.00
Ile	ATA	2.00	7.46	0.17
Ile	ATT	6.00	22.39	0.50
Ile	ATC	4.00	14.93	0.33
Thr	ACG	2.00	7.46	0.29
Thr	ACA	1.00	3.73	0.14
Thr	ACT	1.00	3.73	0.14
Thr	ACC	3.00	11.19	0.43
Trp	TGG	4.00	14.93	1.00
End	TGA	7.00	26.12	0.78
Cys	TGT	2.00	7.46	0.50
Cys	TGC	2.00	7.46	0.50
End	TAG	1.00	3.73	0.11
End	TAA	1.00	3.73	0.11
Tyr	TAT	4.00	14.93	0.50
Tyr	TAC	4.00	14.93	0.50
Leu	TTG	0.00	0.00	0.00
Leu	TTA	0.00	0.00	0.00
Phe	TTT	2.00	7.46	0.33
Phe	TTC	4.00	14.93	0.67
Ser	TCG	5.00	18.66	0.23
Ser	TCA	2.00	7.46	0.09
Ser	TCT	4.00	14.93	0.18
Ser	TCC	6.00	22.39	0.27
Arg	CGG	5.00	18.66	0.17
Arg	CGA	6.00	22.39	0.20
Arg	CGT	8.00	29.85	0.27
Arg	CGC	7.00	26.12	0.23
Gln	CAG	7.00	26.12	0.47
Gln	CAA	8.00	29.85	0.53
His	CAT	4.00	14.93	0.40
His	CAC	6.00	22.39	0.60
Leu	CTG	4.00	14.93	0.36
Leu	CTA	0.00	0.00	0.00
Leu	CTT	3.00	11.19	0.27

Leu	CTC	4.00	14.93	0.36
Pro	CCG	4.00	14.93	0.36
Pro	CCA	2.00	7.46	0.18
Pro	CCT	3.00	11.19	0.27
Pro	CCC	2.00	7.46	0.18

Codon Usage Table for the *Aspergillus parasiticus ver1* gene.

AmAcid	Codon	Number	/1000	Fraction
Gly	GGG	4.00	15.21	0.15
Gly	GGA	6.00	22.81	0.23
Gly	GGT	7.00	26.62	0.27
Gly	GGC	9.00	34.22	0.35
Glu	GAG	11.00	41.83	0.73
Glu	GAA	4.00	15.21	0.27
Asp	GAT	9.00	34.22	0.53
Asp	GAC	8.00	30.42	0.47
Val	GTG	15.00	57.03	0.48
Val	GTA	6.00	22.81	0.19
Val	GTT	3.00	11.41	0.10
Val	GTC	7.00	26.62	0.23
Ala	GCG	6.00	22.81	0.18
Ala	GCA	2.00	7.60	0.06
Ala	GCT	10.00	38.02	0.29
Ala	GCC	16.00	60.84	0.47
Arg	AGG	1.00	3.80	0.06
Arg	AGA	0.00	0.00	0.00
Ser	AGT	1.00	3.80	0.07
Ser	AGC	2.00	7.60	0.14
Lys	AAG	6.00	22.81	0.46
Lys	AAA	7.00	26.62	0.54
Asn	AAT	4.00	15.21	0.44
Asn	AAC	5.00	19.01	0.56
Met	ATG	5.00	19.01	1.00
Ile	ATA	0.00	0.00	0.00
Ile	ATT	5.00	19.01	0.31
Ile	ATC	11.00	41.83	0.69
Thr	ACG	1.00	3.80	0.07
Thr	ACA	3.00	11.41	0.21
Thr	ACT	3.00	11.41	0.21
Thr	ACC	7.00	26.62	0.50
Trp	TGG	2.00	7.60	1.00
End	TGA	0.00	0.00	0.00
Cys	TGT	1.00	3.80	0.25
Cys	TGC	3.00	11.41	0.75
End	TAG	0.00	0.00	0.00
End	TAA	1.00	3.80	1.00
Tyr	TAT	3.00	11.41	0.60
Tyr	TAC	2.00	7.60	0.40
Leu	TTG	3.00	11.41	0.25
Leu	TTA	2.00	7.60	0.17
Phe	TTT	5.00	19.01	0.45
Phe	TTC	6.00	22.81	0.55
Ser	TCG	4.00	15.21	0.29
Ser	TCA	1.00	3.80	0.07
Ser	TCT	3.00	11.41	0.21
Ser	TCC	3.00	11.41	0.21
Arg	CGG	6.00	22.81	0.35
Arg	CGA	1.00	3.80	0.06
Arg	CGT	2.00	7.60	0.12
Arg	CGC	7.00	26.62	0.41
Gln	CAG	4.00	15.21	1.00
Gln	CAA	0.00	0.00	0.00
His	CAT	3.00	11.41	0.50
His	CAC	3.00	11.41	0.50
Leu	CTG	4.00	15.21	0.33
Leu	CTA	0.00	0.00	0.00
Leu	CTT	1.00	3.80	0.08
Leu	CTC	2.00	7.60	0.17

Pro	CCG	0.00	0.00	0.00
Pro	CCA	1.00	3.80	0.14
Pro	CCT	4.00	15.21	0.57
Pro	CCC	2.00	7.60	0.29

Appendix 3.0 Taxonomic Classification of Fungi used in 5.8S rDNA Phylogenetic Analysis.

Aspergillus flavus (Aflav)

Fungi; Ascomycota; Euascomycetes; Plectomycetes; Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

Aspergillus parasiticus (Apara)

Fungi; Ascomycota; Euascomycetes; Plectomycetes; Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

Aspergillus niger (Aniger)

Fungi; Ascomycota; Euascomycetes; Plectomycetes; Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

Emericella nidulans (Anid)

Fungi; Ascomycota; Euascomycetes; Plectomycetes; Eurotiales; Trichocomaceae; Emericella.

Magnaporthe grisea (Mgris)

Fungi; Ascomycota; Euascomycetes; Pyrenomycetes; Phyllachorales; Phyllachoraceae; Magnaporthe.

Colletotrichum graminicola (Cgram)

Fungi; Ascomycota; Euascomycetes; Pyrenomycetes; Phyllachorales; Phyllachoraceae; mitosporic Phyllachoraceae; Colletotrichum.

Neurospora crassa (Ncras)

Fungi; Ascomycota; Euascomycetes; Pyrenomycetes; Sordariales; Sordariaceae; Neurospora.

Alternaria alternata (Aalt)

Fungi; Ascomycota; Euascomycetes; Loculoascomycetes; mitosporic Loculoascomycetes; Alternaria.

Mycosphaerella graminicola (Mgram)

Fungi; Ascomycota; Euascomycetes; Loculoascomycetes; Dothideales; Dothideaceae; Mycosphaerella.

Pyrenophora teres (Pteres)

Fungi; Ascomycota; Euascomycetes; Loculoascomycetes; Dothideales; Pleosporaceae; Pyrenophora.

Mycosphaerella/Dothistroma pini (Dpini)

Fungi; Ascomycota; Euascomycetes; Loculoascomycetes; Dothideales; Dothideaceae; Mycosphaerella.

Filobasidium uniguttulatum (Funig)

Fungi; Basidiomycota; Hymenomycetes; Tremellales; Filobasidiaceae; Filobasidium.

Filobasidiella neoformans (Fneo)

Fungi; Basidiomycota; Hymenomycetes; Tremellales; Filobasidiaceae; Filobasidiella.

Cystofilobasidium capitatum (Ccap)

Fungi; Basidiomycota; Hymenomycetes; Tremellales; Filobasidiaceae; Cystofilobasidium.

REFERENCES.

Agnan, J., Korch, C., and Selitrennikoff, C. 1997. Cloning heterologous genes: problems and approaches. **Fungal Genetics and Biology. 21:** 292-301.

Anderson, J.A., and Green, L.D. 1994. Timing of appearance of versiconal hemiacetal acetate esterase and versiconal cyclase activity in cultures of *Aspergillus parasiticus*. **Mycopathologia. 126:** 169-172.

Assante, G., Locci, R., Camarada, L., Merlini, L., and Nasini, G. 1977. Screening of the genus *Cercospora* for secondary metabolites. **Phytochemistry. 16:** 243-247.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Smith, J.A., Seidman, J.G., and Struhl, K. (eds). 1994. **Current Protocols in Molecular Biology**. John Wiley and Son, New York.

Axley, M.J., Kenigsberg, P., and Hager, L.P. 1986. Fructose induces and glucose represses chloroperoxidase mRNA levels. **The Journal of Biological Chemistry. 32:** 15058-15061.

Ballance, D.J. 1986. Sequences important for gene expression in filamentous fungi. **Yeast. 2:** 229-236.

Bandelt, H., and Dress, A.W.M. 1992. Split Decomposition: a new and useful approach to phylogenetic analysis of distance data. **Molecular Phylogenetics and Evolution. 1:** 242-252.

Bassett, C., and Buchanan, M. 1970. A toxic Difuroanthraquinone from *Dothistroma pini*. **Chemistry and Industry.** 26 December 1970**:** 1659-1660.

Bidlake, A. T. 1996. Development of an efficient transformation system for *Dothistroma pini*. MSc Thesis. Massey University, New Zealand.

Blanke, S.R., and Hager, L.P. 1988. Identification of the fifth axial heme ligand of chloroperoxidase. **The Journal of Biological Chemistry. 263:** 18739-18743.

Blanke, S.R., and Hager, L.P. 1990. Chemical modification of chloroperoxidase with diethylpyrocarbonate. **The Journal of Biological Chemistry. 265:** 12454-12461.

Bowden, C.G., Hintz, W.E., Jeng, R., Hubbes, M., and Horgen, P.A. 1994. Isolation and characterisation of the cerato-ulmin toxin gene of the Dutch elm disease pathogen, *Ophiostoma ulmi*. **Current Genetics 25:** 323-329.

Bowden, C.G., Smalley, E., Guries, R.P., Hubbes, M., Temple, B., and Horgen P.A. 1996. Lack of association between cerato-ulmin production and virulence in *Ophiostoma novo-ulmi*. **Molecular Plant-Microbe Interactions 9:** 556-564.

Bradshaw, R.E., Bennett, J.W., and Peberdy, J.F. 1983. Parasexual analysis of *Aspergillus parasiticus*. **Journal of General Microbiology. 129:** 2117-2123.

Brown, D., Yu, W. J., Kelkar, H.S., Fernandes, M., Nesbitt, T.C., Keller, N.P., Adams, T.H., and Leonard, T.J. 1996. Twenty-five coregulated transcripts define a sterigmatocystin gene cluster in *Aspergillus nidulans*. **Proceedings of the National Academy of Sciences of the United States of America. 93:** 1418-1422.

Bruchez, J.J.P., Eberle, J., and Russo, V.E.A. 1993. Regulatory sequences involved in the translation of *Neurospora crassa* mRNA: Kozak sequences and stop codons. **Fungal Genetics Newsletter. 40:** 85-88.

Buades, C., and Moya, A. 1996. Phylogenetic analysis of the isopenicillin-N-synthetase horizontal gene transfer. **Journal of Molecular Evolution. 42:** 537-542.

Bullock, W.O., Fernandez, J.M., and Short, J.M. 1987. XL-1-Blue: a high efficiency plasmid transforming *recA Escherichia coli* strain with β-galactosidase selection. **Biotechniques. 5:** 376-378.

Burow, G.B., Nesbitt, T.C., Dunlap, J., and Keller, N.P. 1997. Seed lipoxygenase products modulate *Aspergillus* mycotoxin biosynthesis. **Molecular Plant-Microbe Interactions. 10:** 380-387.

Burrows, L.L., Charter, D.F., and Lam, J.S. 1996. Molecular characterisation of the *Pseudomonas aeruginosa* serotype O5 (PAO1) B-band lipopolysaccharide gene cluster. **Molecular Microbiology. 22:** 481-495.

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*). **Current Genetics. 18:** 347-354.

Carson, S.D., and Carson, M.J. 1989. Breeding for resistance in forest trees, a quantitative genetic approach. **Annual Review of Phytopathology 27:** 373-395.

Carson, S.D., and Carson, M.J., 1991. Realising gains in resistance to *Dothistroma*. **Proceedings of the 11th Research Working Group meeting (Forest Genetics).** 11-15 March, Coonwarra, South Australia.

Cary, F.W., Wright, M., Bhatnagar, D., Lee, R., and Chu, F.S. 1996. Molecular characterisation of an *Aspergillus parasiticus* dehydrogenase gene, *norA*, located on the aflatoxin biosynthesis gene cluster. **Applied and Environmental Microbiology. 62:** 360-366.

Chang, P., Skory, C.D., and Linz, J.E. 1992. Cloning of a gene associated with aflatoxin B_1 biosynthesis in *Aspergillus parasiticus*. **Current Genetics. 21:** 231-233.

Chang, P., Cary, J.W., Bhatnagar, D., Cleveland, T.E., Bennett, J.W., Linz, J.E., Woloshuk, C.P., and Payne, G.A. 1993. Cloning of the *Aspergillus parasiticus apa-2* gene associated with the regulation of aflatoxin biosynthesis. **Applied and Environmental Microbiology. 59:** 3273-3279.

Chang, P., Ehrlich, K., Yu, J., Bhatnagar, D., and Cleveland, T.E. 1995. Increased expression of *Aspergillus parasiticus aflR*, encoding a sequence-specific DNA-binding

protein, relieves nitrate inhibition of aflatoxin biosynthesis. **Applied and Environmental Microbiology. 61:** 2372-2377.

Chen, Z., Jiang, J.C., Lin, Z., Lee, W.R., Baker, M.E. and Chang, S.H. 1993. Site-specific mutagenesis of *Drosophila* alcohol dehydrogenase: evidence for involvement of Tyr152 and Lys156 in catalysis. **Biochemistry. 32:** 3342-3346.

Chumley, F.G., and Valent, B. 1990. Genetic analysis of melanin-deficient, non-pathogenic mutants of *Magnaporthe grisea*. **Molecular Plant-Microbe Interactions. 3:** 135-143.

Cleveland, T.E., Lax, A.R., Lee, L.S., and Bhatnagar, D. 1987. Appearance of enzyme activities catalyzing conversion of sterigmatocystin to aflatoxin B_1 in late-growth-phase *Aspergillus parasiticus* cultures. **Applied and Environmental Microbiology. 53:** 1711-1713.

Cohen, S.N., Chang, A.C.Y., and Hsu, L. 1972. Non-chromosomal antibiotic resistance in bacteria: Genetic transformation of *E. coli* by R-factor DNA. **Proceedings of the National Academy of Sciences of the United States of America. 69:** 2110.

Danks, A.V., and Hodges, R. 1974. Polyhydroxyanthraquinones from *Dothistroma pini*. **Australian Journal of Chemistry. 27:** 1603-1606.

Dick, A.M.P. 1989. Control of *Dothistroma* needle blight in the *Pinus radiata* stands of Kinleith forest. **New Zealand Journal of Forestry Science. 19:** 171-179.

Elliott, G.S., Mason, R.W., Ferry, D.G., and Edwards, I.R. 1989. Dothistromin risk assessment for forestry workers. **New Zealand Journal of Forestry Science. 19:** 163-170.

Evans, H.C. 1984. **The genus *Mycosphararella* and its anamorphs *Cercoseptoria,Dothistroma* and *Lecanosticta* on pines**. Commonwealth Agricultural Bureaux, Surrey.

Feng, G.H., and Leonard, T.J. 1998. Culture conditions control expression of the genes for aflatoxin and sterigmatocystin biosynthesis in *Aspergillus parasiticus* and *A. nidulans*. **Applied and Environmental Microbiology. 64:** 2275-2277.

Flaherty J.E., Weaver, M.A., Payne, G.A., and Woloshuk, C.P. 1995. A β-glucuronidase reporter gene construct for monitoring aflatoxin biosynthesis in *Aspergillus flavus*. **Applied and Environmental Microbiology. 61:** 2482-2486.

Flaherty, J.E., and Payne, G.A. 1997. Overexpression of *aflR* leads to upregulation of pathway gene transcription and increased aflatoxin production in *Aspergillus flavus*. **Applied and Environmental Microbiology. 63:** 3995-4000.

Franich, R.A. and Wells, L.G. 1977. Infection of *Pinus radiata* by *Dothistroma pini*: effect of buffer capacity of needle homogenates. **New Zealand Journal of Forestry Science. 7:** 35-39.

Franich, R.A., Wells, L.G., and Barnett, J.R. 1977. Variation with tree age of needle cuticle topography and stomatal structure in *Pinus radiata* D. Don. **Ann. Bot. 41:** 621-626.

Franich, R.A., Wells, L.G., and Holland, P.T. 1978. Epicuticular wax of *Pinus radiata* needles. **Phytochemistry 17:** 1617-1623.

Franich, R.A. 1981. Determination of dothistromin by quantitative reversed-phase thin-layer chromatography. **Journal of Chromatography. 209:** 117-120.

Franich, R.A. 1988. Chemistry of weathering and solubilisation of copper fungicide and the effect of copper on germination, growth, metabolism, and reproduction of *Dothistroma pini*. **New Zealand Journal of Forestry Science 18:** 318-328.

Gadgil, P.D. 1967. Infection of *Pinus radiata* needles by *Dothistroma pini*. **New Zealand Journal of Botany 5:** 498-503.

Gadgil, P.D., and Hilden, G. 1976. Effect of light intensity on infection of *Pinus radiata* by *Dothistroma pini*. **New Zealand Journal of forestry science 6:** 67-71.

Gadgil, P.D. 1984. *Dothistroma* needle blight. **Forest Pathology in New Zealand 5:** 1-8.

Gallagher, R.T., and Hodges, R. 1972. The chemistry of dothistromin, a difluroanthraquinone from *Dothistroma pini*. **Australian Journal of Chemistry. 25:** 2399-2407.

Gillman, C.J. 1996. Identification of dothistromin biosynthetic pathway genes. MSc Thesis. Massey University, New Zealand.

Guzmán-de-Peña, D., and Ruiz-Herrera, J. 1997. Relationship between aflatoxin biosynthesis and sporulation in *Aspergillus parasiticus*. **Fungal Genetics and Biology. 21:** 198-205.

Hamer, J.E., and Timberlake, W.E. 1987. Functional organisation of the *Aspergillus nidulans trpC* promoter. **Molecular and Cellular Biology. 7:** 2352-2359.

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. **Journal of Molecular Biology. 166:** 557-580.

Harvey, A.M., Batt, R.D., and Pritchard, G.G. 1976. Inhibition of RNA synthesis in *Chlorella pyrenoidosa* and *Bacillus megaterium* by the pine-blight toxin, dothistromin. **Journal of General Microbiology. 95:**268-276.

Hibbett, D.S. 1996. Phylogenetic evidence for horizontal transmission of group I introns in the nuclear ribosomal DNA of mushroom-forming fungi. **Molecular Biology and Evolution. 13:** 903-917.

Hicks, J.K., Yu, J., Keller, N.P., and Adams, T.H. 1997. *Aspergillus* sporulation and mycotoxin production both require inactivation of the FadA Gα protein-dependent signaling pathway. **The EMBO Journal. 16:** 4916-4923.

Hirst, P. 1997. Genetic diversity of *Dothistroma pini* in New Zealand. MSc Thesis. Massey University, New Zealand.

Holmes, D.S., and Quigley, M. 1981. A rapid boiling method for the preparation of bacterial plasmids. **Analytical Biochemistry. 114:** 193-197.

Hopwood, D.A. 1997. Genetic contributions to understanding polyketide synthases. **Chemical Reviews. 97:** 2465-2497.

Huson, D.H. 1998. SplitsTree: analyzing and visualizing evolutionary data. **Bioinformatics. 14:** 68-73.

Itakura, K., Rossi, J.J., and Wallace, R.B. 1994. Synthesis and use of synthetic oligonucleotides. **Annual Review of Biochemistry. 53:** 323.

Jones, W.T., Harvey, D., Jones, S.D., Fielder, S., Debnam, P., and Reynolds, P.H.S. 1993. Competitive ELISA employing monoclonal antibodies specific for dothistromin. **Food and Agricultural Immunology. 5:** 187-197.

Jones, W.T., Harvey, D., Sutherland, P.W., and Reynolds, P.H.S. 1998. Production of anti-idiotypic monoclonal antibodies that mimic the phytotoxin dothistromin. **Food and Agricultural Immunology. 10:** 67-78.

Kawamura, C., Moriwaki, J., Kimura, N., Fujita, Y., Fuji, S., Hirano, T., Koizumi, S., and Tsuge, T. 1997. The melanin biosynthesis genes of *Alternaria alternata* can restore pathogenicity of the melanin-deficient mutants of *Magnaporthe grisea*. **Molecular Plant-Microbe Interactions. 10:** 446-453.

Kelkar, H.S., Keller, N.P., and Adams, T.H. 1996. *Aspergillus nidulans stcP* encodes an *O*-methyltransferase that is required for sterigmatocystin biosynthesis. **Applied and Environmental Microbiology**. 62: 4296-4298.

Kelkar, H.S., Skloss, T.W., Haw, J.F., Keller, N.P., and Adams, T.H. 1997. *Aspergillus nidulans stcL* encodes a putative cytochrome P-450 monooxygenase required for bisfuran desaturation during aflatoxin/sterigmatocystin biosynthesis. **The Journal of Biological Chemistry. 272:** 1589-1594.

Keller, N.P., Kantz, N.J., Adams, T.H. 1994. *Aspergillus nidulans verA* is required for the production of the mycotoxin sterigmatocystin. **Applied and Environmental Microbiology. 60:** 1444-1450.

Keller, N.P., Segner, S., Bhatnagar, D., and Adams, T.H. 1995. *stcS*, a putative P-450 monooxygenase, is required for the conversion of versicolorin A to sterigmatocystin in *Aspergillus nidulans*. **Applied and Environmental Microbiology. 61:** 3628-3632.

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

Kempken, F. 1995. Horizontal transfer of a mitochondrial plasmid. **Molecular General Genetics. 248:** 89-94

Kimura, N., and Tsuge, T. 1993. Gene cluster involved in melanin biosynthesis of the filamentous fungus *Alternaria alternata*. **Journal of Bacteriology. 175:** 4427-4435.

Kiser, K.B., and Lee, J.C. 1998. *Staphylococcus aureus cap5O* and *cap5P* genes functionally complement mutations affecting enterobacterial common-antigen biosynthesis in *Escherichia coli*. **Journal of Bacteriology. 180:** 403-406.

Klich, M.A., Montalbano, B., and Ehrlich, K. 1997. Northern analysis of aflatoxin biosynthetic genes in *Aspergillus parasiticus* and *Aspergillus sojae*. **Applied Microbiology and Biotechnology. 47**: 246-249.

Kubo, Y., Takano, Y., and Furusawa, I. 1996. Molecular genetic analysis of the melanin biosynthetic genes essential for appressorium function in *Colletotrichum lagenarium*. **Molecular Aspects of Pathogenicity: Requirements for Signal Transduction** (D. Mills, H. Kunoh, N. Keen, and S. Maryama, Eds.) pages 73-82. APS Press, St. Paul, USA.

Kusumoto, K., Mori, K., Nogata, Y., Ohta, H., and Manabe, M. 1996. Homologs of the aflatoxin biosynthetic gene *ver-1* in strains of *Aspergillus oryzae* and related species. **Journal of Fermentation and Bioengineering. 82:** 161-164.

Lan, R., and Reeves, P.R. 1996. Gene transfer is a major factor in bacterial evolution. **Molecular Biology and Evolution. 13:** 47-55.

Lawrence, J.G., and Roth, J.R. 1996. Selfish operons: horizontal transfer may drive the evolution of gene clusters. **Genetics. 143:** 1843-1860.

Lawrence, J.G. 1997. Selfish operons and speciation by gene transfer. **Trends in Microbiology. 5:** 355-359.

Lee, L. S., Bennett, J.W., Cuculla, A.F., and Stanley, J.B. 1975. Synthesis of versicolorin A by a mutant strain of *Aspergillus parasiticus* deficient in aflatoxin production. **Journal of Agricultural and Food Chemistry. 23:** 1132-1134.

Li, W. 1997. **Molecular Evolution.** Sinauer Associates, Sunderland.

Liang, S., Skory, C.D., and Linz, J.E. 1996. Characterisation of the *ver-1A* and *ver-1B* genes involved in aflatoxin biosynthesis in *Aspergillus parasiticus*. **Applied and Environmental Microbiology. 62:** 4568-4575.

Liang, S., Wu, T., Lee, R., Chu, F.S., and Linz, J.E. 1997. Analysis of mechanisms regulating expression of the *ver1* gene, involved in aflatoxin biosynthesis. **Applied and Environmental Microbiology. 63:** 1058-1065.

Lin, B., and Anderson, J.A. 1992. Purification and properties of versiconal cyclase from *Aspergillus parasiticus*. **Archives of Biochemistry and Biophysics. 293:** 67-70.

Liu, S.L., and Dean, R.A. 1997. G protein α subunit genes control growth, development, and pathogenicity of *Magnaporthe grisea*. **Molecular Plant-Microbe Interactions. 10:** 1075-1086.

Masel, A.M., He, C., Poplawski, A.M., Irwin, J.A.G., and Manners, J.M. 1996. Molecular evidence for chromosome transfer between biotypes of *Colletotrichum gloeosporioides*. **Molecular Plant-Microbe Interactions. 9:** 339-348.

Matsushima, K., Ando, Y., Hamasaki, T., and Yabe, K. 1994. Purification and characterisation of two versiconal hemiacetal acetate reductases involved in aflatoxin biosynthesis. **Applied and Environmental Microbiology. 60:** 2561-2567.

Messing, J. 1987. New M13 vectors for cloning. **Methods in Enzymology. 101:** 20-73.

Minto, R.E., and Townsend, C.A. 1997. Enzymology and molecular biology of aflatoxin biosynthesis. **Chemical Reviews. 97:** 2537-2555.

Morgan, B. 1998. Isolation of a polyketide synthase gene from *Dothistroma pini*. MSc Thesis. Massey University, New Zealand.

Nakamura, Y., Gojobori, T., and Ikemura, T. 1998. Codon usage tabulated from the international DNA sequence database. **Nucleic Acids Research. 26:** 334.

Nuell, M.J., Fang, G., Axley, M.J., Kenigsberg, P., and Hager, L.P. 1988. Isolation and nucleotide sequence of the chloroperoxidase gene from *Caldariomyces fumago*. **Journal of Bacteriology. 170:** 1007-1011.

Panaccione, D.G., Scott-Craig, J.S., Pocard, J., and Walton, J.D. 1992. A cyclic peptide synthase gene required for pathogenicity of the fungus *Cochliobolus carbonum* on maize. **Proceedings of the National Academy of Sciences of the United States of America. 89:** 6590-6594.

Payne, G.A., Nystrom, G.J., Bhatnagar, D., Cleveland, T.E., and Woloshuk, C.P. 1993. Cloning of the *afl-2* gene involved in aflatoxin biosynthesis from *Aspergillus flavus*. **Applied and Environmental Microbiology. 59:** 156-162.

Perpetua, N.S., Kubo, Y., Yasuda, N., Takano, Y., and Furusawa, I. 1996. Cloning and characterisation of a melanin biosynthetic *THR1* reductase gene essential for appressorial penetration of *Colletotrichum lagenarium*. **Molecular Plant-Microbe Interactions. 9:** 323-329.

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

Prade, R.A., Griffith, J., Kochut, K., Arnold, J., and Timberlake, W.E. 1997. *In vitro* reconstruction of the *Aspergillus* (=*Emericella*) *nidulans* genome. **Proceedings of the National Academy of Sciences of the United States of America. 94:** 14564-14569.

Prieto, R., Yousibova, G.L., and Woloshuk, C.P. 1996. Identification of aflatoxin biosynthesis genes by genetic complementation in an *Aspergillus flavus* mutant lacking the aflatoxin gene cluster. **Applied and Environmental Microbiology. 62:** 3567-3571.

Prieto, R., and Woloshuk, C.P. 1997. *ordl*, an oxidoreductase gene responsible for the conversion of O-methylsterigmatocystin to aflatoxin in *Aspergillus flavus*. **Applied and Environmental Microbiology. 63:** 1661-1666.

Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. **Molecular Cloning: A Laboratory Manual.** New York: Cold Spring Harbour Laboratory Press.

Sanger, F., Nicklen, S., and Coulson, A.R. 1977. DNA sequencing with chain terminating inhibitors. **Proceedings of the National Academy of Sciences of the United States of America. 74:** 5463-5467.

Shain, L., and Franich, R.A. 1981. Induction of *Dothistroma* blight symptoms with dothistromin. **Physiological Plant Pathology. 19:** 49-55.

Shaw, G.J., Chick, M. and Hodges, R. 1978. A ^{13}C NMR Study of the biosynthesis of the anthraquinone dothistromin by *Dothistroma pini*. **Phytochemistry. 17:** 1743-1745.

Shimizu, K., Tanaka, C., and Tsuda, M. 1997. Cloning of *Brn1*, a reductase gene involved in melanin biosynthesis in *Cochliobolus heterostrophus*. **Journal of General and Applied Microbiology. 43:** 145-150.

Silva, J.C., and Townsend, C.A. 1996. Heterologous expression, isolation, and characterisation of versicolorin B synthase from *Aspergillus parasiticus*. **The Journal of Biological Chemistry. 272:** 804-813.

Skory, C.D., Chang, P.K., Cary, J., and Linz, J.E. 1992. Isolation and characterisation of a gene from *Aspergillus parasiticus* associated with the conversion of versicolorin A to sterigmatocystin in aflatoxin biosynthesis. **Applied and Environmental Microbiology. 58:** 3527-37.

Skory, C.D., Chang, P., and Linz, J.E. 1993. Regulated expression of the *nor-1* and *ver-1* genes associated with aflatoxin biosynthesis. **Applied and Environmental Microbiology. 59:** 1642-1646.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. **Journal of Molecular Biology. 98:** 503-517.

Steyn, P.S (ed). 1980. **The Biosynthesis of Mycotoxins - A Study in Secondary Metabolism**. Academic Press.

Stoessl, A. 1984. Dothistromin as a metabolite of *Cercospora arachidicola*. **Mycopathologia. 86:** 165-168.

Stoessl, A., Abramowski, Z., Lester, H.H., Rock, G.L., and Towers, G. 1990. Further toxic properties of the fungal metabolite dothistromin. **Mycopathologia 112:** 179-186.

Storck, R., and Alexopoulos, C.J. 1970. Deoxyribonucleic acid of fungi. **Bacteriology Reviews. 34:** 126-154.

Takano, Y., Kubo, Y., Shimiizu, K., Mise, K., Okuno, T., and Furusawa, I. 1995. Structural analysis of *PKS1*, a polyketide synthase gene involved in melanin biosynthesis in *Colletotrichum lagenarium*. **Molecular General Genetics. 249:** 162-167.

Takano, Y., Kubo, Y., Kawamura, C., Tsuge, T., and Furusawa, I. 1997. The *Alternaria alternata* melanin biosynthesis gene restores appressorial melanization and penetration of cellulose membranes in the melanin-deficient albino mutant of *Colletotrichum lagenarium*. **Fungal Genetics and Biology. 21:** 131-140.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. 1994. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting,

position-specific gap penalties and weight matrix choice. **Nucleic Acids Research. 22:** 4673-4680.

Trail, F., Mahanti, N., and Linz, J. 1995. Molecular biology of aflatoxin biosynthesis. **Microbiology. 141:** 755-765.

Upchurch, R.G., Walker, D.C., Rollins, J.A., Ehrenshaft, M., and Daub, M.E. 1991. Mutants of *Cercospora kikuchii* altered in cercosporin synthesis and pathogenicity. **Applied and Environmental Microbiology. 57:** 2940-2945.

van den Brink, H.J.M., van Gorcom, R.F.M., van den Hondel, C.A.M.J.J., and Punt, P.J. 1998. Cytochrome P-450 enzyme systems in Fungi. **Fungal Genetics and Biology. 23:** 1-17.

Vidal-Cros, A., Viviani, F., Labesse, G., Boccara, M., and Gaudry, M. 1994. Polyhydroxynaphthalene reductase involved in melanin biosynthesis in *Magnaporthe grisea*. **European Journal of Biochemistry. 219:** 985-992.

Wang, X. 1997. Genetic variability in the canker pathogen fungus, *Gremmeniella abietina*. Contribution of sexual compared with asexual reproduction. **Mycological Research. 101:** 1195-1201.

Weigel, B.J., Burgett, S.G., Chen, V.J., Skatrud, P.L., Frolik, C.A., Queener, S.W., and Ingolia, T.D. 1988. Cloning and expression in *Escherichia coli* of isopenicillin-N-synthetase genes from *Streptomyces lipmanii* and *Aspergillus nidulans*. **Journal of Bacteriology. 170:** 3817-3826.

White, T.J., Bruns, T.D., Lee, S.B., and Taylor, J.W. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenies. In **PCR protocols: A Guide to Methods and Applications** (Innis, Gelfand, Sninsky, and White, Eds). pages 315-322. Academic Press, New York.

Woloshuk, C.P., Foutz, K.R., Brewer, J.F., Bhatnagar, D., Cleveland, T.E., and Payne, G.A. 1994. Molecular characterisation of *aflR*, a regulatory locus for aflatoxin biosynthesis. **Applied and Environmental Microbiology. 60:** 2408-2414.

Woloshuk, C.P., Yousibova, G.L., Rollins, J.A., Bhatnagar, D., and Payne, G.A. 1995. Molecular characterisation of the *afl-1* locus in *Aspergillus flavus*. **Applied and Environmental Microbiology. 61:** 3019-3023.

Woloshuk, C.P., and Prieto, R. 1998. Genetic organisation and function of the aflatoxin B_1 biosynthetic genes. **FEMS Microbiology Letters. 160:** 169-176.

Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. **Gene. 33:** 103-119.

Yu, J., Cary, J.W., Bhatnagar, D., Cleveland, T.E., Keller, N.P., and Chu, F.S. 1993. Cloning and characterisation of a cDNA from *Aspergillus parasiticus* encoding an *O*-methyltransferase involved in aflatoxin biosynthesis. **Applied and Environmental Microbiology. 59:** 3564-3571.

Yu, J., and Leonard, T.J. 1995. Sterigmatocystin biosynthesis in *Aspergillus nidulans* requires a novel type I polyketide synthase. **Journal of Bacteriology. 177:** 4792-4800.

Yu, J., Chang, P.K., Cary, J.W., Wright, M., Bhatnagar, D., Cleveland, T.E., Payne, G.A., and Linz, J.E. 1995. Comparative mapping of aflatoxin pathway gene clusters in *Aspergillus parasiticus* and *Aspergillus flavus*. **Applied and Environmental Microbiology. 61:** 2365-2371.