CHARACTERISATION OF A PUTATIVE DOTHISTROMIN BIOSYNTHETIC CLUSTER.

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.

Seth Laarakkers

1999

ABSTRACT.

The fungus *Dothistroma pini* is a key pathogen in New Zealand (and international) softwood plantations, most notably *P. radiata*. The mycotoxin dothistromin produced by this saprophytic fungus is believed to play a major role in its pathogenesis.

Dothistromin shares functional groups and pathway intermediates with those of sterigmatocystin and aflatoxin, secondary metabolites of *Aspergillus* sp. As the sterigmatocystin and aflatoxin biosynthetic pathways are characterised this provided us with a model pathway and potential probes for the isolation of dothistromin genes.

The *ver*1 gene is critical to the completion of aflatoxin biosynthesis in *Aspergillus* sp. as its disruption prevented the synthesis of aflatoxin. Assuming similar enzymes act in the dothistromin biosynthetic pathway a probe for *ver*1 was obtained and used to probe a *D. pini* genomic library. This led to the isolation of two lambda clones named λCGV1 and λCGV2 (Gillman 1996). A second library screen was completed using an aflatoxin polyketide synthase (PKS) probe and led to the isolation of the lambda clone λBMKSA (Morgan 1997).

The λCGV1 clone has been studied in detail and shown to contain a gene similar to aflatoxin *ver*1 (named *dkr*1) and other potential dothistromin biosynthetic genes (Monahan 1998). This study looks in greater detail at the lambda clones λCGV2 and λBMKSA and determines whether they contain putative dothistromin biosynthetic genes and are part of the anticipated gene cluster.

In this project the lambda clone λCGV2 was partially characterised which revealed that the other potential *ver* gene showed a greater similarity to the melanin biosynthetic gene *phn* than to the aflatoxin gene *ver*-1. This implied that the clone was unlikely to contain dothistromin biosynthetic genes so no further sequence was generated. However, a partial restriction map was constructed.

The other lambda clone, λBMKSA was then further characterised. Double stranded sequence of the putative *pks* gene region was completed. The remainder of the lambda clone was subcloned and exploratory sequence revealed a gene with high similarity to *stc*W.

The next stage was to determine how the three lambda clones were related. This was approached by probing genomic Southern blots with the ends of the lambda clones to determine the presence of commonly hybridised fragments. The presence of common fragments suggests that the three clones are very close together in the genome, although the evidence which links λCGV2 and λBMKSA is stronger than the evidence that links λCGV2 and λCGV1.

This is the first evidence that the three lambda clones isolated using aflatoxin probes are close together in the genome of *D. pini.* The genes present on these lambda clones show a high degree of similarity to their aflatoxin counterparts and could potentially contain a dothistriomin biosynthetic cluster.

Acknowledgements.

Firstly I must thank both my supervisor Rosie Bradshaw, without your encouragement and guidance this project would never have been completed, and my family for their on going support and understanding.

Thanks to everybody at Massey, in and out of the MGU, for answering my many questions and being good company. All of the suggestions, discussions (arguments!) and explanations made for my two most interesting years at Massey. The best of luck to all of you in your degrees, research and future careers.

Thanks to all of my friends away from Massey, Jim without your help I don't know what I would have done, your friendship is invaluable. Everybody else, thanks for the company, diverse conversations, and consistent laughter.

Thanks to Jaime, Eru, Tala, and all the other people I trained with for giving me time I was to busy to think about Massey! Thanks to Chuck and Benjamin for trips to remind me that NZ has the best bush and coastline in the world for us to enjoy, and providing contrast to the Massey environment, thanks also to Rhys and the teams at Flyers and Icons for all the laughter and team spirit they provided, it was always a pleasure to work with you.

Finally, thanks to Stewart for scanning all my Autorads, and everybody else who helped assemble this thesis.

TABLE OF CONTENTS.

2. MATERIALS AND METHODS.

LIST OF TABLES.

LIST OF FIGURES.

1. INTRODUCTION.

1.1 New Zealand Exotic Forest Damage.

New Zealand is quite fortunate to be free of many internationally significant exotic diseases in its forestry. The most damaging factor to forestry in this country is wind, followed, in order of decreasing significance, by pathogens, fire, erosion and lastly frost and snow damage. Of all exotic forestry pathogens one of the most damaging is *Dothistroma pini*. Investigation into the costs of *D. pini* infection in the North Island reveal that the annual direct cost of controlling this infection is 1.6 million NZ dollars. Further, the predicted residual growth losses are approximately 4.5 million NZ dollars annually. It is obvious that even if by international standards *D. pini* is insignificant, *D. pini* is a most significant pathogen in New Zealand (New 1989, Ministry of Forestry 1996).

1.1.1 *Dothistroma pini.*

D. pini is the saprophytic fungus responsible for Dothistroma needle blight, most notably in *Pinus* sp. In this condition the host is infected by multinucleate spores which are dispersed mainly by raindrop splashes but also by other vectors such as wind (Sheridan 1970). These spores germinate and lead to the formation of characteristic brick red necrotic bands on the needles and premature defoliation. This ultimately leads to reduced photosynthetic ability and, eventually, the death of the host (Gadgil 1984).

D. pini (Hulbary) is a filamentous fungus which occurs in both sexually and asexually reproducing forms. Taxonomically the imperfect *D. pini* is of the order Coleomyctes and

class Deuteromycotina. The sexually reproducing form is referred to as *Scirrhia pini* (also *Mycospharella pini*) and is of the order Dothideales and class Ascomycotina (Gadgil 1967).

D. pini was reported in New Zealand in 1962 (Chou 1991, Dick 1989). Prior to this, it was identified in large plantations in Tanzania in 1957 (Gibson 1972). This disease may have entered New Zealand through numerous vectors, ranging from the trouser turnups of a person who visited an infected forest, to long range spore dispersal by natural means.

To date *D. pini* infections have been reported from all parts of the North Island and from the West Coast and Invercargill in the South Island of New Zealand (Gadgil 1984). As well as the New Zealand infections *D. pini* is well established world wide. From where it was first a threat (Tanzania) it spread to all major plantations in Kenya, Uganda, Malawi and Zimbabwe by 1964. Dothistroma spread through England to the rest of Europe, but was only a major pathogen in Spain (Gibson 1972). By the 1960s *D. pini* was well established in North, central and South America. In Asia *D. pini* has been reported from India and Japan, and in 1975 it appeared in Australia (Ministry of Forestry, F.R.I. Bulletin No.18). More recently *D. pini* has been observed at in Portugal and Germany (Proceedings of IUFRO 1997).

1.2 Infection.

One of the key factors determining the severity of *D. pini* infection is the duration of the period of high moisture immediately after spore dispersal (humidity). Germination of hydrated conidia occurred at a range of temperature and moisture combinations tested, but the only combination to produce a high rate of infection included constant moisture at 24^{o}C

for day duration and 16°C for the night duration (Sheridan 1970, Gadgil 1974). Further, it was shown that severity of infection also depends heavily on the length of the dry period following infection. The longer the dry period, the less severe the infection (Gadgil 1976). Moisture is also important to stages of the *D. pini* life cycle other than germination, for example spore release and dispersal, which only occur when there is a high amount of surrounding moisture.

Light intensity does not affect spore germination in *D. pini* and it has no effect on the early growth of *D. pini* on the needle surface. However, the severity of infection increased linearly with decreasing light intensity, this is also due to the effect of decreasing light intensity on the host plant. This is apparent in the observation that shaded foliage was markedly less infected than that foliage which was fully exposed to light. This observation is in part also due to the effect of light on other factors involved such as the toxic *D. pini* secondary metabolite dothistromin which will also be present (Gadgil 1976).

In an early study of seasonal spore presence it was shown that season had no direct effect on presence of the spores as *D. pini* stromata on infected needles produce viable conidia throughout the year. Tests on the duration of spore survival demonstrated that the spores survived longer, up to six months, on suspended needles than in leaf litter where they all died in 2-4 months. This means that when the spores contact the ground they are not as great a source of infection as when first released from suspended needles. The spores in the leaf litter died much more quickly when the leaf litter was dry. Pruning and thinning trees can further reduce the incidence of reinfection from leaf litter as this increases airflow and leads to less moisture in the leaf litter. This is true for all microbial activity in the leaf litter (Gadgil 1969).

In Africa (Kenya) Gibson showed that conidia could be taken up into clouds by evaporation and in the low temperatures within the clouds they could easily survive. Early conidial survival tests showed that conidia could survive for up to eleven months in dry pine needles at 18^{o}C and five months at 30^{o}C (Gibson 1964). According to a more recent report this estimate of spore life span is generous, but there is still a potential for long-distance spread (Gadgil 1969).

1.2.1 Host invasion and Colonisation.

Forty-eight hours after conidia contact the needle they germinate and form one to three germ tubes per conidium. Only approximately ten percent of conidia fail to germinate. Next, for a period of seven to ten days, profuse mycelial growth occurs on the needle surface. By this stage (about ten days later) appressorium like structures form on needles (usually over stomatal openings) and secondary conidia are produced. Mycelial growth does not continue to persist on the needle surface. Instead of this, secondary conidia provide an ongoing source of infection. The hyphae grow from the mycelia and 'search' for openings, possibly following a pH gradient as observed in *Uromyces vicae-fabae* (Edwards 1986). When a hypha encounters an opening, usually a stomatal pore, it will grow through it to enter the inner cells of the needle. It is worth noting that direct access through the epidermis is observed only when needles are inoculated with macerated mycelia. Once the fungus has gained access to the interior of the needle hyphal growth continues both inter and intracellularly causing disruption of the mesophyll (Gadgil 1967). It is at this stage that dothistromin is released. Dothistromin is believed to disrupt the cells of the mesophyll tissue surrounding the site of infection, but well in advance of the hyphae, allowing the

saprophytic *D. pini* to extract nutrients from the local dead mesophyll tissue (Gadgil 1984). Then groups of small black stromata penetrate the epidermis and, upon contact with water, swell and burst to release multinucleate spores. The water then runs over foliage, and falls to the ground. The spores are dispersed by raindrop splashes. It follows that infection is primarily from neighbour to neighbour (Sheridan 1970).

Once the host tree has been infected, one response of *P. radiata* to needle mesophyll tissue damage by dothistromin is to synthesise benzoic acid in the surrounding tissue. Benzoic acid is toxic to needle mesophyll and fungistatic to *D. pini.* This is believed to be the first observed example of a conifer phytoalexin response. In this case the supposed phytoalexin is benzoic acid and it causes the death of cells adjacent to those which are infected. This in turn leads to the eventual death of infected needles. This supports the observation that long lesion length is related to a stronger resistance in the host, as a faster, stronger benzoic acid response would lead to larger lesions (Franich 1986). However, these lesion length observations have not been supported in further observations (Carson Per. Comm.).

1.3 Dothistroma Blight.

Infection usually begins at the base of the crown of the host tree. This infection leads to the formation of the characteristic brick red bands on the needles which persist long after needle necrosis. These bands are distinctly separate regions to the rest of the infected needle and eventually small black dots appear on them which are the fruiting bodies (stromata) which release new conidia (Gadgil 1984). The red pigment is the toxic fungal secondary metabolite dothistromin. The primary infection is followed by local cell necrosis which leads to the death of the needle and eventually premature defoliation of the area, related to a

loss of photosynthetic ability, and possibly the death of the host (Shaw 1977). Dothistroma blight most readily affects softwood species such as *Pinus attentuata, P. nigra, P. ponderosa* and *P. jeffreyi,* however from an economic viewpoint in New Zealand the most important species infected is *Pinus radiata,* especially in young stands. New Zealand natives grow relatively slowly (estimated mean annual increment increase of 1-2 cubic metres per hectare per year) compared to *P. radiata* (estimated mean annual increment increase of 20-25 cubic metres per hectare per year). *P. radiata* is also less ecologically sensitive so more easily maintained. This makes *P. radiata* the obvious species of choice to support the New Zealand pulp and paper industry, as well as supply construction materials to local and export markets. New Zealand has a large exotic forestry, 90.5% of which is made up of *P. radiata* trees worth a potential 5.5 billion NZ dollars by 2010. The sum of New Zealand's forestry is spread over 1.5 million hectares of land (Ministry of Forestry 1996). No other pathogen in New Zealand has been in the position to cause so much damage to the potential forestry earnings and related business in this country.

In areas where *D. pini* is left unchecked over a five year period, approximately six square metres of basal area is lost per hectare. (The basal area decrease refers to the loss in covered ground area due to mortality and growth losses in the diameter increment of trees.) This equates to 24% of the total basal area. Also 1.6m of height (or 13%), over five years, is lost from green crown height. This is believed to be an underestimate as Whyte showed a 3.9m loss in green crown height in a comparable experiment (Woolons 1984).

Other studies have shown a 17-73% loss of annual growth in younger trees. An important factor to consider from an economic viewpoint is that there is little visible impact until

25+% defoliation has occurred in more than 50% of the stand (Shaw 1977, Ministry of Forestry, F.R.I. Bulletin No.18).

1.4 Control of *D. pini* Infection.

1.4.1 Quarantine.

Recent studies of the *D. pini* population by RAPD analysis have shown that New Zealand has only one genetic type of *D. pini*, possibly due to all New Zealand infections stemming from a single spore (Hirst 1997). This lack of genetic diversity in all New Zealand isolates, coupled with the possibility of long range spread of *D. pini* spores via rain clouds (Gadgil 1984), forces us to consider the implications of the introduction of a different strain of *D. pini* on the forestry of New Zealand. If another strain of *D. pini*, of a different mating type to the strain already present, entered the *D. pini* population of New Zealand there would be the opportunity for sexual recombination. At present New Zealand is fortunate enough to have very few *P. radiata* pathogens of any significance. This is probably due to a combination of good luck and the quarantine system developed in the 1940's under which all wood imports are inspected (including sawn and unsawn logs, cases, pallets, and other wood based packaging). Any evidence of the presence of insects or fungi leads to the treatment of the entire container. Also, the importation of bark is prohibited due to the greater likelihood of bark boring insects or other pathogens being present and going unnoticed (due to the degree of difficulty involved in the inspection of large volumes of bark) (NZ Customs pers. Comm.). As well as all countries infected with *D. pini*, the primary countries from which imports are of concern are North America and Europe as these temperate countries are more likely to hold pathogens that can be problematic in New

Zealand than central Asian countries where the tropical pathogens are probably unable to survive well enough to be of any great consequence (Bulman 1992). As the life cycle of *D. pini* is considerably shorter than that of its host it is likely that in the event of the introduction of a second, sexual strain, the recombination and diversity developed could lead to the evolution of a more virulent or fungicide resistant mutant which could devastate New Zealand's exotic forests. To avoid this potential disaster we should examine the quarantine system to ensure that it excludes all possible sources of *D. pini* spores.

Due to the primitive nature of present *D. pini* control techniques (see below) it is important that current research uncovers more information about this pathogen and the manner in which it functions, in order to mitigate its potentially disastrous effects on forestry.

1.4.2 Spraying.

To date the only practical method of control of *D. pini* has been the aerial application of copper-based fungicides. This practice has been common in New Zealand for over 20 years. Both cuprous oxide and copper oxychloride have been used to equal effect, but since 1972 copper oxychloride has been preferred due to its lower cost. New Zealand uses up to 700 tonnes of copper-based fungicide a year. At present the fungicide is applied by aircraft equipped with Micron air spraying equipment which allows accurate control of droplet size and has sizeably reduced the volume of spray required.

The optimum time for spray application varies from year to year but is generally in between October and December for a single spray and between January and February for the second spray if it is required. The disease warrants spraying when the infected crown level reaches

15-25% for a single spray but a double spray if it is higher than this. If the spray dries before rain it is unnecessary to reapply (Assessment and control of *Dothistroma* needle blight, F.R.I. Bulletin). In the central North Island of New Zealand a total of 300mm of rain falls about every three months. This is enough to remove most of the applied copper, but reinfection is slow and respraying is not required for about three years (Gadgil, 1969). The exposure of spores to 20mg of copper ions per litre of spray kills them in 1.5 hours. On *P. radiata* needles copper fungicide forms copper complexes with surface molecules giving more water-soluble copper.

An application rate of 30 mg/l of copper ions is sufficient to reduce *D. pini* germination rates to less than pathogenic levels, and even 10 mg/l copper ions is sufficient to reduce germ tube length and reduce the number of hyphal penetrations, and 5 mg/l is enough to reduce the production of secondary conidia by surface mycelia. However lower concentrations of 1-5 mg/l increase the rate of secondary metabolism so increase the production of dothistromin (Franich, 1988).

An exception to this general statement applies in nurseries. Due to the age of these pines they are very susceptible to high infection levels and, as nurseries supply wide areas with seedlings, They provide possible source of primary infection to new, previously uninfected areas. *P. radiata* are close stocked in nursery beds, which also encourages disease and can lead to high infection levels. When seedlings reach 2.5cm in height they should be sprayed monthly with 4kg of copper oxychloride per hectare (Assessment and control of *Dothistroma* needle blight, F.R.I. Bulletin no.18).

1.4.3 Host Resistance in Mature Trees.

Some hosts (e.g. Pinus *radiata*) are highly susceptible to infection until about ten years of age when they develop a high resistance. Electron microscope studies of the changes in pine needles as the tree ages revealed several major structural variations, the most important of these to *D. pini* are a decreased stomatal pore size from 15-20 micrometers to 10-15 micrometers (Franich 1977). Another key change occurs in the stomatal guard cell epidermis. When the fungus establishes itself through the stomata in a young tree the guard cell epidermis is covered with a fine microtubular wax (nonacosan-10-ol and long chain alkane diols). When the hypha contacts this wax it expands to form an appressorium structure and hyphae penetrate between the guard cells. However in older, *D. pini* resistant trees stomatal regions have been shown to be occluded with amphorous wax. This wax prevents hyphal contact with the fine microtubular wax contacted in young plants. This may act as a mechanical barrier to prevent hyphal contact, obscure a chemotactic signal required by hyphae, or even contain compounds which inhibit hyphal growth. Chemical purification and investigation into the wax has shown it to contain an inhibitor of *D. pini* spore germination but further studies are required (Franich 1982, Franich 1983).

1.4.3.1 Selective Breeding for Host Resistance.

D. pini resistance in *P. radiata* has been shown to be due to the actions of genes at several loci. This is referred to as a quantitative genetic trait. This makes conventional breeding difficult as increased selection for certain traits may unknowingly reduce the selection for other desirable traits which are less phenotypically obvious (Carson 1989). Fortunately selection for stem straightness and low branching is not affected by selection for *D. pini*

resistance, however the genetic gain in growth rate is. This means that it is favourable to plant resistant trees in areas where there is expected to be a high incidence of *D. pini* infection, but other types of trees will produce a greater yield in other areas (Carson 1991).

In 1983 a disease resistance breeding programme was initiated. The trees produced by this programme have shown a 7%-20% reduction in infection by *D. pini* infection (Carson 1989). The most advanced tree breeding programmes are only into about their third generation due to the long life cycle of *P. radiata*. This has led to many difficulties in performing classical genetic experiments. However with the advent of molecular biological techniques soon the genetic characteristics of the trees may be determined at embryo level so the entire process could be accelerated (Moffat, 1996).

As these advanced new breeds of tree are developed and tested it is expected that they will reduce the average infection level in *P. radiata* by approximately 15% in areas of medium to high infection. This should be cumulative with the effects of spraying, but spraying cannot reduce mean infection below 5%. For these reasons it would be favourable to devise a new, more effective and economically viable means of control of *D. pini* (Carson 1991).

A prime New Zealand example of the cost of a *D. pini* infection is Kinleith forest. In 1966 the *D. pini* infection of this forest was addressed by the aerial application of copper fungicide (copper oxychloride). It was decided that if 15-30% of the crowns in a stand were infected they would be sprayed in early summer, and if 30+% were infected they would be sprayed once early, and once late in summer. Up to 1988 the cost of this was

18.4 million NZ dollars. The costs in later years were less due to improvements in spray application techniques leading to a reduction in total spray volume.

If a breed of *D. pini* resistant trees had been planted at Kinleith instead of the standard breed, it is estimated that savings of up to 56% (NZ$467,000) would have been made annually on spray costs alone (Dick 1989).

1.4.3.2 Clonal Forestry.

Clonal forestry is steadily gaining popularity. This aspect of forestry is likely to grow in importance as molecular techniques are used to introduce genes to host trees which will then be propagated clonally. In the current traditional breeding system 15-60 unrelated parents are open pollinated and a wide variety of individuals with a broad genetic base are produced. However in true clonal forestry large forests can be planted of a single clone, this obviously leads to a loss of genomic diversity and gives disease the opportunity to become catastrophic. There is also the possibility of a genetic sampling error, such as an incorrect belief that a certain phenotype is relevant to an underlying characteristic, leading to an increased susceptibility to infection. Frequently throughout history diseases which have never posed a significant threat have grown to catastrophic proportions when the host population was made more genetically uniform by basic agriculture. Although clonal forestry may give us the opportunity to introduce resistance genes to a particular disease we must be aware of the possibility of this making other pathogens more severe (Carson, F.R.I. Rotorua).

In New Zealand there have been large scale monoculture plantings for over sixty years (Chou 1991). The foresters have always been concerned with diseases and they have had outbreaks of several damaging pine diseases, fortunately all were treatable. One of the pathogens was *Colletotrichum acutatum*, a fungus which until then had never been classified as a pathogen. This inevitably draws the question of whether this was due to the monocultured nature of the forestry. There have been no truly catastrophic losses in New Zealand's clonal forestry, however practices such as 75-85% thinning by 10-12 years old coupled with short rotation and clear felling (25-30 year rotation) considerably reduce the obvious effects of pathogens (Chou 1991).

The introduction of resistant and more expensive seedlings may increase the value of the trees to the point where this degree of thinning is financially unacceptable, and in doing so make the control of damage due to pathogens more important.

1.5 Host Fungi Interaction.

Fungi are a group of highly versatile eukaryotic carbon heterotrophic organisms the majority of which are obligate saprophytes. Only about 10% of over 100 000 known species are capable of causing disease, yet more than 300 000 species of flowering plant exist and all are affected by pathogenic fungi. A single host can be colonised by only a few species of fungi, and all exant fungi have a limited range so tend to be specialised to an individual plant species. These phytopathogenic fungi cause persistent annual crop losses and are capable of causing devastating epidemics (Heath 1996).

Underlying the infection of a host by a fungal pathogen are many complex mechanisms and the genotypes of both the host and parasite determine the outcome of infection. Environmental stresses, such as infection by pathogens, stimulate many local and systemic reactions. An important point to note is that the defence response of a plant is an allelic inheritable trait rather than an adaptive trait as observed in mammals (Heath 1996).

Extensive study of the interactions between *Linum usitatissimum* (flax) and the obligate flax rust pathogen *Melampsora lini* lead Flor to propose that the outcome of plant-fungus interaction was determined by a gene for gene interaction. This is now referred to as the elicitor-receptor model and is the mechanism by which a resistance gene product in the host encodes a receptor that binds a pathogen avirulence gene product leading to early pathogen recognition. Once the plant has recognised the presence of a pathogen the primary response is the generation of highly active oxygen species such as hydrogen peroxide in the vicinity of the pathogen to prevent its spread (oxidative burst). This elicitor induces defence reactions in all non-host species but not in host species (Heath 1996).

A compatibility factor subverts the plants defence response to allow colonisation of the host by the fungi. Compatibility factors are often recognised as host selective toxins. Compatibility factors are typically toxic at high concentrations. It is difficult to distinguish whether the *D. pini* secondary metabolite dothistromin is an elicitor, a compatibility factor, or simply a toxin.

1.6 Dothistromin.

Primary metabolic compounds and the enzymes required for their production and modification perform basic, well-defined and often essential roles in the support of growth, repair, and reproduction of an organism. However, secondary metabolites are compounds not obviously required for the basic processes of growth and reproduction, but may function to enhance an organisms ability to fulfil its ecological niche (Drew 1977).

This important secondary metabolite of *D. pini* is visible as the brick red pigment which is characteristic of Dothistroma blight and is also produced by other species such as the peanut pathogen *Cercospora arachidicola* (Liang 1996). Dothistromin was shown to inhibit RNA synthesis using the test organisms *Chlorella pynenoidosa* and *Bacillus megaterium,* but its phytotoxicity is to be established (Stoessl 1984). The inhibition of RNA synthesis is due to the inhibition of the incorporation of uridine into RNA thus inhibiting growth. This inhibition occurs with 1-2μM of dothistromin and had effects similar to actinomycin D. This is quite different to aflatoxin which inhibits RNA synthesis through extensively binding to DNA.

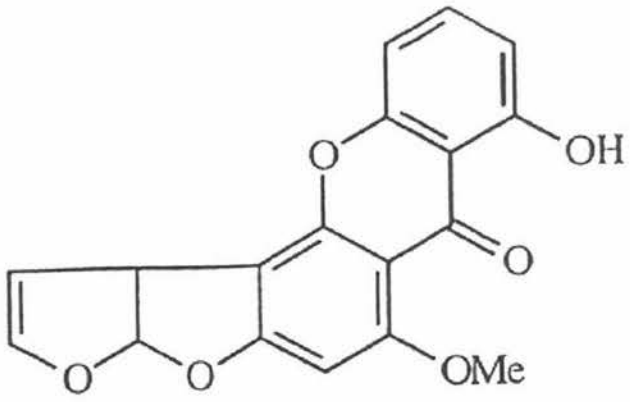
When purified dothistromin 10-100ng (in acetone) is inoculated into *P. radiata* needles it induces all of the symptoms of Dothistroma blight. As treatment with acetone alone caused no symptoms, it follows that dothistromin plays a key role in the pathogenesis of *D. pini* (Shain 1981). Natural lesions contain 1-10μg of dothistromin which is substantially more than is required to induce symptoms in the laboratory. The induced symptoms developed within hours of exposure to dothistromin and continued to develop for the entire seventy

two hour test period. The symptoms appeared more quickly in high light intensity. In the dark there was a significant increase in the levels of ethylene produced. These results suggest that tissue necrosis occurs only when there is an interaction between dothistromin and photosynthetically active tissue (Shain 1981). The role of dothistromin in the disease processes suggests its control may carry potential for a means of controlling *D. pini*.

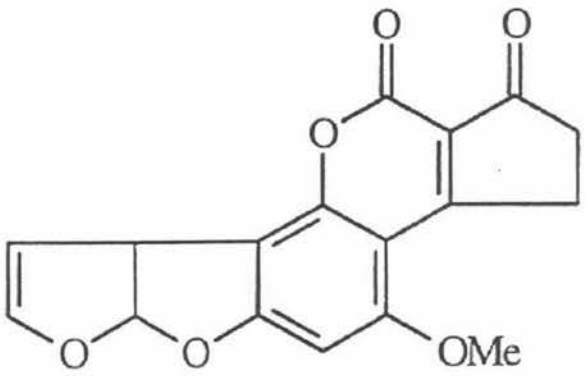
1.6.1 Chemistry of Dothistromin.

Dothistromin ($C_{18}H_{12}O_9$) is a secondary fungal metabolite possessing a furobenzofuro moiety considered to be of importance to the functional properties of aflatoxin (Gallagher 1972). This places dothistromin in a category containing many potent carcinogens and mutagens. Other members of this group include the sterigmatocystin and versicolorin groups, as well as the aflatoxins produced by certain *Aspergillus* species (Figure 1.1). Dothistromin is the first anthraquinone in this series of toxins. As aflatoxin is also a secondary fungal metabolite and one of the most potent carcinogens ever tested, it followed that dothistromin required testing of its mutagenic and carcinogenic properties (Gallagher 1972, Kelly 1973).

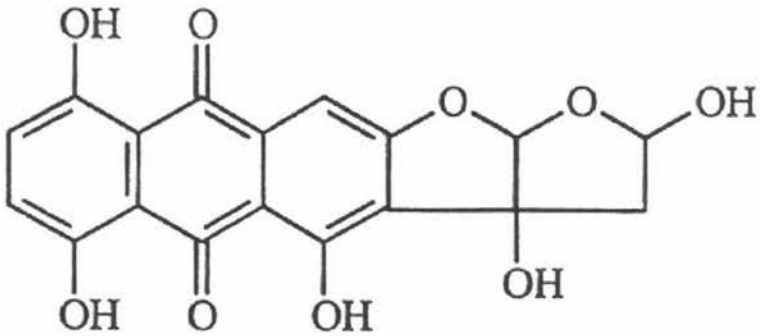
Figure 1.1 Structures of sterigmatocystin, aflatoxin B1, and dothistromin.



Sterigmatocystin



Aflatoxin B1



Dothistromin

Baker showed dothistromin to be positive for three mutagenicity assays (sister chromatid exchange, micronucleus formation, and the Ames test) which shows that irrespective of metabolic activation dothistromin shows an increase in chromosomal damage (Baker 1989).

Further testing for toxicity showed it to be ten times more active than aflatoxin B1 (Most powerful mutagen tested to date) at inhibiting growth of *B.megaterium,* and in 1986 dothistromin, or its metabolite, was shown to cause clastogenic, recombinogenic and point mutational effects in mammalian cells (Ferguson 1986).

1.7 Gene Clusters and Their Regulation.

In studies of the production of secondary metabolites of fungi it is apparent that they are synthesised in multistep pathways, the precursors of which are often present as intermediates of the primary metabolism of that fungi. The enzymes required are often of broad substrate specificity. In filamentous fungi these genes are sometimes clustered and may contain as many as twenty five genes in regions of up to 60kb of DNA for a single pathway (Keller 1997).

An example of a characterised fungal biosynthetic pathway can be found in the sterigmatocystin gene cluster in *A. nidulans.* This cluster is estimated to include genes which code at least 15 enzymatic functions, all contained in a 60kb region of DNA. This gene clustering may be important for the correct regulation of the biosynthetic pathway as the appearance of an enzyme before that of its correct substrate could prevent the formation of the final product (Brown 1996).

The regulation of these biosynthetic pathways is proposed to involve a control method in which a regulatory protein is essential. This is probably also true of some species such as *Streptomyces griseus* which are inhibited or repressed by the presence of glucose or easily utilisable phosphate, carbon, or nitrogen suggesting that the promoters of most secondary metabolites are activated only under specific environmental conditions (Martin 1989).

A well documented example of fungal secondary metabolite production regulation is the aflatoxin biosynthetic pathway. In this pathway the maximum transcription of two key genes (*Nor-1* and *Ver-1*) occurs just prior to the onset of the stationary phase of growth (idiophase). Approximately eight hours later the presence of aflatoxin B1 is observed (Skory 1992). A fungal mutant complementation analysis in *A. parasiticus* identified the gene *Apa*-2 as being able to restore a wild type phenotype on mutants unable to produce aflatoxin. Over expression of this gene resulted in the overproduction of all of the pathway intermediates. These two points led to the conclusion that *Apa*-2 is a regulatory gene for the production of aflatoxin. Further analysis showed that the predicted amino acid sequence of this gene contains a cysteine rich region which is homologous to the C6 subgroups of many *S. cerevisiae* regulatory genes (Chang 1993).

Afl R was identified in *A. flavus* as regulating aflatoxin biosynthesis. Sequence analysis has since revealed that it is nearly identical to the *A. parasiticus Apa*-2 gene. The sequence analysis of these suspected regulatory sequences showed that they encode a putative Zinc finger associated with DNA binding proteins. This further suggests that it is a possible regulatory protein associated with gene expression (Chang 1995).

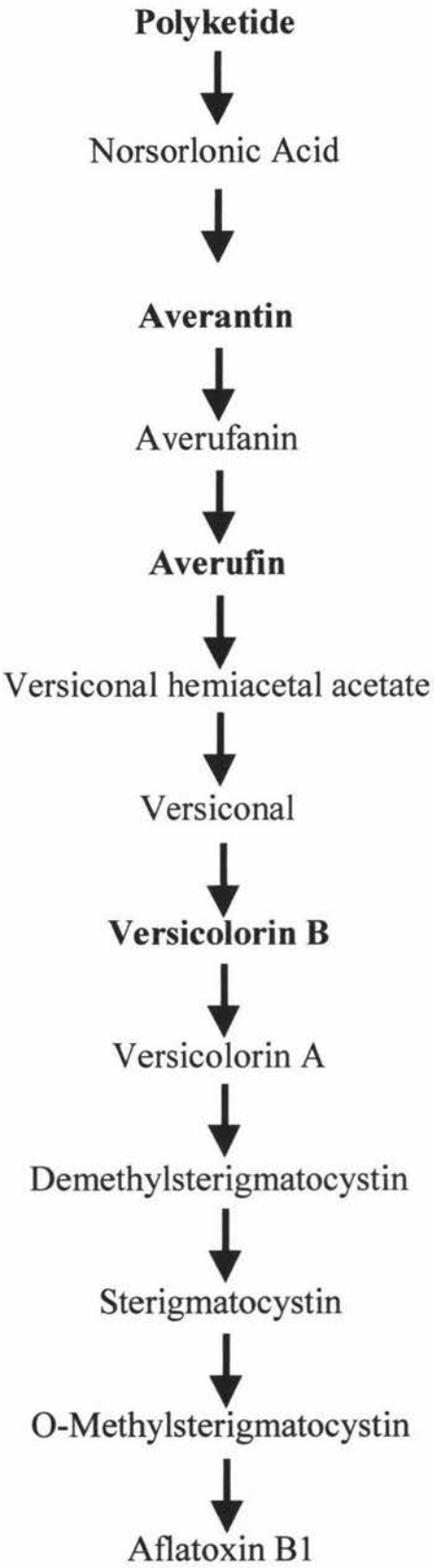
1.8 Aflatoxin and Sterigmatocystin.

Aflatoxin and sterigmatocystin are carcinogenic metabolites of the same biochemical pathway. The *A. nidulans* sterigmatocystin cluster contains 25 distinct genes labelled *stc* followed by the letter of the alphabet assigned by position in the cluster.

Aflatoxin biosynthesis in some *Aspergillus* species shows similarities to dothistromin biosynthesis (Figure 1.2). As *Aspergillus* species are also filamentous fungi, it follows that the methods employed to isolate genes in aflatoxin biosynthetic pathway are probably applicable to the dothistromin biosynthetic pathway. Aflatoxin is synthesised by the condensation of acetate units. The order of the genes in the clusters in *A. parasiticus* and *A. flavus* is the same, but the order of *A. nidulans* is different. However, as the map of the biosynthetic pathway genes of *A. parasiticus* and *A. flavus* is not complete, this may change. The spacing between the genes in this cluster shows a low degree of variability. The genes themselves are highly conserved between the species.

Figure 1.2 Comparison of aflatoxin biosynthetic pathway with dothistromin biosynthetic pathway.

Common intermediates highlighted in bold.


Polyketide
Norsorlonic Acid
Averantin
Averufanin
Averufin
Versiconal hemiacetal acetate
Versiconal
Versicolorin B
Versicolorin A
Demethylsterigmatocystin
Sterigmatocystin
O-Methylsterigmatocystin
Aflatoxin B1

1.9 *Ver* Genes

1.9.1 Aspergillus *Ver* Genes.

A gene associated with the conversion of versicolorin A to sterigmatocystin in the aflatoxin biosynthetic pathway, *ver*1A, was first isolated from *A. parasiticus.* This was obtained by complementation of an aflatoxin minus mutant with a DNA fragment from a wild type strain (Skory 1992). When this gene was used as a probe on a genomic Southern blot of *A. parasiticus* another region of homology was detected. This second gene was isolated and named *ver*1B (Keller 1994).

Further investigation demonstrated *ver*1B was 95% identical to *ver*1A. Upon sequencing it was determined that *ver*1B contained a nonsense mutation which led to the formation of a stop codon in its sequence. This gene coded only a truncated polypeptide with no apparent function. This observation was confirmed by the use of a linearised plasmid to knockout the *ver*1A gene which lead to the accumulation of the pathway intermediates at this step in the aflatoxin pathway thus proving that the *ver*1B gene had no relevant function.

Further research revealed that two copies of the *ver* gene were also present in *Aspergillus nidulans,* but as yet only one copy has been isolated in Aspergillus *flavus* (Silva 1996). In *A. nidulans* the two *ver* genes are both functional and are believed to be involved in one step or two adjacent steps in the pathway.

In *A. nidulans* the *stc*S (*ver*1B) disruption was unable to produce sterigmatocystin, the pathway was blocked at versicolorin A. Further characterisation of *stc*S showed that it encoded a cytochrome P-450 monooxygenase. It is unclear how many enzymatic reactions are required per step in the pathway. Some steps may require several enzymes (as seen in the *A. nidulans* versicolorin-aflatoxin conversion) and some enzymes may help catalyse several steps (Keller 1995).

Recently it has been determined that a chromosomal region of about 12kb was duplicated in *Aspergillus parasiticus,* and this region also contained the proposed regulatory gene in the cluster, *Apa*-2 (Liang 1996). The levels of both dothistromin and aflatoxin produced by some of the wild type strains tends to be more than is required to induce symptoms. This is possibly due to the overproduction of a gene product because the regulatory gene is present in the duplicated region. This adds evidence to the possibility that *Apa*-2 is one of the regulatory genes in this cluster (Wolfe 1997).

Recent studies of the yeast genome show that there are 376 pairs of homologous genes which show 23-100% sequence identity present in *S. cerevisiae.* Often these genes are phenotypically redundant, but 13% of yeast proteins are believed to be derived from duplications. This duplication of large sections, or even the entire genome are believed to be a major evolutionary mechanism (Wolfe 1997). This is similar to the observed area of duplication in *A. parasiticus.*

The *A. parasiticus* amino acid sequence of the *ver*1A gene is 66% identical and 82% similar to a tetrahydroxynapthalene reductase (T_4HN reductase) in *Magnaporthe grisea.* This reductase reduces emodin to chrysophanol at an early stage of fungal pigment biosynthesis

mediated by NADPH (Vidal-Cross 1994). The examination of the T_4HN reductase in *Magnaporthe grisea* suggests it is capable of reduction of both trihydroxynapthalene (T_3HN) and T_4HN so is responsible for two metabolic reductive steps. These reactions are similar to the versicolorin A to sterigmatocystin conversion in aflatoxin biosynthesis. The high degree of sequence identity between *ver*1 and T_4HN reductase suggests that the protein coded by *ver*1 is the counterpart of T_4HN reductase for the aflatoxin biosynthetic pathway. This observation is in favour of the hypothesis that versicolorin A is processed in two successive steps.

1.9.2 *Dothistroma pini ver* Genes.

Comparisons between the *ver*-1 gene of *D. pini* (*dkr*1) and those of some *Aspergillus* species have shown a high degree of similarity. An *A. parasiticus verA* gene was used to probe a Southern blot of the *D. pini* genome. Initially this probe hybridised to four bands on the blot at low stringency. The regions of highest intensity corresponded to the sizes of bands of the λCGV1 library clone which had been isolated using the *A. parasiticus verA* gene as a probe. Part of this clone was then sequenced and compared to the GenBank database where it predicted a 78% sequence identity to both the *A. parasiticus* and the *A. nidulans ver*-1, as well as a 62% predicted identity to a *M. grisea* gene (Gillman, 1996). This first *ver* gene to be isolated from *D. pini* was named *ver*1. Since then a second clone has been identified by the same probe. This clone is named *ver*2.

1.10 Altered Expression Mutants in Fungi.

The soybean pathogen *Cercospora kikuchii* produces a light activated compound, cercosporin, which is important in the development of disease symptoms. The disease has similar symptoms to Dothistroma blight. Mutants which have lost the ability to produce the toxin cercosporin have been isolated and studies conducted on these mutants revealed that they were non pathogenic (Upchurch, 1991). This means that there is a potential for using molecular biological techniques to produce mutants which do not produce toxins. Unfortunately, co-fermentation studies of two different mutants leading to the aflatoxin minus phenotype revealed that aflatoxin was produced. This is because aflatoxin precursors were secreted into the medium by one strain and taken up by the other, this is referred to as cross feeding of pathway intermediates. There was a tendency for more of the metabolites near the end of the pathway to be released and the levels of aflatoxin produced were relatively low compared to the wild type levels of production. One other key factor is that high moisture may be required for cross feeding to occur (Cleveland 1991). This leads to the conclusion that if gene knockouts are to be used to attempt to biologically control the levels of aflatoxin produced it may be a better plan to target genes early in the pathway or genes required for the regulation of secondary metabolite biosynthesis.

1.11 Objectives of This Project.

This project will focus on three aspects of the putative dothistromin biosynthetic genes:

- Sequencing the potential *ver*2 gene in the λCGV2 clone.
- Sequencing the potential *pks* gene in the λBMKSA clone.
- Assessing how the three lambda clones fit together.

The lambda clone λCGV2 was believed to contain a second copy of the putative dothistromin biosynthetic gene *ver* (Gillman1996). As one of the aims of the dothistromin programme is to disrupt the *ver*1 gene to produce a dothistromin minus mutant and assess its pathogenicity, the possibility of gene duplication must be addressed. The potential second copy of *ver* required characterisation to determine if it could potentially substitute a disrupted *ver*1 gene and the aim of this project was to perform a sequence comparison. The relevance of this is exemplified in *Ophiostoma ulmi* where the production of an apparent pathogenicity factor, cerato-ulmin, was disrupted with no effect on pathogenicity. Before any further research targeting dothistromin as a means of *D. pini* control can be carried out the relevance of dothistromin must be clearly demonstrated.

The lambda clone λBMKSA was previously isolated and partial sequence obtained (Morgan 1998). The putative polyketide synthase gene required double stranded sequencing to remove anomalies. The lambda clone λBMKSA also required comparison to the previously characterised lambda clones λCGV1 and λCGV2 to determine whether they overlapped. Exploratory sequence of this clone would also be generated to determine if other genes relevant to dothistromin biosynthesis were contained.

The three characterised lambda clones: λCGVI, λCGV2 and λBMKSA, all shown to contain potential dothistromin biosynthetic genes were believed to be clustered as gene clustering was present in the model aflatoxin biosynthetic pathway. The relations between the three lambda clones needed to be determined to assess whether these genes were part of a potential dothistromin biosynthetic cluster.

2. MATERIALS AND METHODS.

2.1 BACTERIAL AND FUNGAL STRAINS, LAMBDA CLONES, AND PLASMIDS.

The bacterial and fungal 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 purified water and sterilised by autoclaving at 15 p.s.i. (121 °C) for 15 minutes. Liquid media were then cooled to room temperature before the addition of supplements. Solid media was cooled to 50 °C before the addition of supplements and pouring. All uninoculated plates were stored at 4 °C.

2.2.1 Luria-Bertaini (LB) Media.

Luria Broth (LB) media contained (g/L): Tryptone (Difco), 10.0; NaCl, 5.0; yeast extract (Oxoid) 5.0; and where referred to as 'LB Maltose' 2.0 g/L of maltose (Oxoid). The media was dissolved and the pH adjusted to 7.4 prior to autoclaving. For solid media, agar (Davis) was added to 15.0 g/L. LB maltose was also supplemented with $MgSO_4.7H_2O$ at a final concentration of 10 mM.

Table 2.1 Bacterial And Fungal Strains, λ Clones and Plasmids..

Strain, λ Clone, Plasmid.	Relevant Features.	Source or reference.
Bacterial Strains:		
Escherichia coli.		
XL-1 Blue	*sup*E44 *hsd*R17 *rec*A1 *end*A1 *gyr*A46 *thi rel*A1 *lac*$^-$ F' [*pro*AB$^+$ *lac*^{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*A1λ-	Hanahan, 1983.
KW251	F$^-$ *sup*E44 *sup*F58 *gal*K2 *gal*T22 *Met*B1 *hsd*R2 *mcr*B1 *mcr*A$^-$ *arg*A81:*Tn*10 *rec*D1014	Murray et al., 1977
Fungal Strains:		
Dp2	Single spore isolate of wild type	Long Mile Road, NZFRI, Rotorua
λ Clones:		
λCGV1	LambdaGEM-12 (Promega) containing genomic DNA (including *ver*1) from Dp2.	Gillman, 1996
λCGV2	LambdaGEM-12 (Promega) containing genomic DNA (including *phn*1) from Dp2.	Gillman, 1996
λBMKSA	LambdaGEM-12 (Promega) containing genomic DNA (including *pks*) from Dp2.	Morgan, 1996

Plasmids:

pUC118	Amp^r *lacZ*'(3.2 kb)	Messing, 1983
R144	pUC118 containing 1.0 kb *Bam*HI from λCGV2.	This study.
R145	pUC118 containing 2.0 kb *Bam*HI/ *Eco*RI from λCGV2.	This study.
R146	pUC118 containing 2.45 kb *Bam*HI from λCGV2.	This study.
R147	pUC118 containing 4.4 kb *Bam*HI from λCGV2.	This study.
R151	pUC8 containing 2.5 kb *Bam*HI/ *Eco*RI from λCGV1.	Monahan, 1998.
R156	pUC8 containing 2.4 kb *Bam*HI/ *Eco*RI from λBMKSA.	Morgan, 1998.
R161	pUC118 containing 3.0 kb *Bam*HI/ *Eco*RI from λBMKSA.	This study.
R162	pUC118 containing 4.5 kb *Bam*HI from λBMKSA.	This study.
R163	pUC118 containing 3.5 kb *Eco*RI from λBMKSA.	This study.
R181	pUC118 containing 0.94 kb *Bam*HI/ *Eco*RI from λBMKSA.	This study.

When required LB was supplemented after autoclaving to give final concentrations of: ampicillin, 100 μg/ml; tetracycline, 10 μg/ml; isopropylthio-β-D-galactoside (IPTG), 30μg/ml; and 3-indolyl-β-D-galactoside (X-gal) at 60 μg/ml.

2.2.2 Terrific Broth.

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

2.2.3 TB Top Agar.

TB top agar contained (g/L): tryptone (Difco), 10.0; NaCl, 5.0; agar (Davis), 8.0. This was autoclaved and then cooled to approximately 50°C before being supplemented to a final concentration of 10 mM $MgSO_4$.

2.2.4 NZCYM.

NZCYM contained (g/L): NZ amine, 10.0; NaCl, 5.0; Casamino acids (Difco), 1.0; yeast extract (Oxoid). 5.0; $MgSO_4.7H_2O$, 2.0. Prior to autoclaving the pH was adjusted to 7.5.

2.2.5 *Dothistroma* Media (DM).

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

2.3 GROWTH AND MAINTENANCE OF CULTURES.

2.3.1 Fungal Cultures.

An 8 mm cube of *D. pini* mycelia was ground in 1 ml of Milli-Q purified water in a 1.5 ml microcentrifuge tube using an autoclaved plastic grinder. 200 µl of this crude mycelium suspension was plated onto DM plates overlaid with cellophane discs. The cultures were sealed with parafilm and grown in the dark at 20 °C for 10-14 days. Cultures were stored at 4 °C for up to 6 months before subculturing.

2.3.2 Bacterial Cultures.

E. coli cultures were grown overnight (approximately 16 hours) either on solid or in liquid media with the appropriate supplements (Section 2.2.1). Plates were sealed with parafilm and stored in the dark at 4 °C for up to 6 weeks before subculturing by streaking onto fresh, supplemented LB plates. For long term storage a 3 ml overnight liquid culture was pelleted by a 1 minute microcentrifuge spin (13 000r.p.m.) and the pellet was resuspended in 1ml of 20% (v/v) glycerol and snap frozen in dry ice before being stored at –70 °C.

2.4 BUFFERS AND SOLUTIONS.

All buffers and solutions were prepared using Milli-Q purified water and sterilised by autoclaving at 15 p.s.i. (121°C) for 15 minutes unless otherwise stated.

2.4.1 TE Buffer.

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.

1 X TAE buffer contained 40 mM Tris-HCl, 2 mM Na_2EDTA and 20 mM acetic acid, pH 8.5.

2.4.3 1 x TBE Buffer.

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 20 x SSC.

20 x SSC contained 3 M NaCl and 0.2 M tri-sodium citrate. This was then diluted to the required concentration with Milli-Q purified water as required.

2.4.5 Ethidium Bromide.

The ethidium bromide used to stain DNA in agarose gels was prepared by the addition of 1 µl of a 10 mg/ml stock per 10 ml of Milli-Q purified water to give a final concentration of 1 µg/ml.

2.4.6 RNase A (DNase free).

RNase A contained 10 mg/ml pancreatic RNase A (Sigma) in 0.01 M Na_2EDTA (pH 5.2). This was then heated to 100 °C for 15 minutes, then allowed to cool slowly to room temperature where the pH was adjusted by the addition of 0.1 volumes of 1 M Tris (pH 7.4). The solution was dispensed into 1 ml aliquots and stored at –20 °C.

2.4.7 STET Buffer.

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).

2.4.8 10 x Gel Loading Buffer (GLB).

10 x GLB 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.

2.4.9 Hybridisation Solution.

Hybridisation solution contained (per litre): 150 ml of 20 x SSC (Section 2.4.4), 20 ml of 50 x Denhardt's (section 2.4.11), 30 ml of 10% (w/v) SDS and 5ml of 50 μg/ml herring sperm DNA. The solution was manufactured from autoclaved ingredients in an autoclaved container and did not require autoclaving.

2.4.10 50 x Denhardt's Solution.

50 x Denhardt's solution contained (g/l): ficoll, 10.0; polyvinylpyrolidone, 10.0; bovine serum albumin, 10.0. The solution was filter sterilised.

2.4.11 10 x TNE Solution

TNE solution contained (g/l): Tris, 121; EDTA-Na_2, 3.7; Nacl, 58.4. This was filter purified and stored at 4 °C.

2.4.12 SM Buffer.

SM Buffer contained (g/l): NaCl, 5.8; $MgSO_4.7H_2O$, 2.0; and 50ml/l 1M Tris-HCl (pH 7.5)

2.5 DNA ISOLATION.

2.5.1 Rapid Boiling Plasmid Preparation.

Plasmid DNA was isolated as described by Holmes and Quigley (1981). The DNA was resuspended after purification in 20 µl of autoclaved Milli-Q purified water or TE (10:0.1; section 2.4.1).

2.5.2 Alkaline Lysis Plasmid Preparation (Small Scale).

This method for isolating plasmid DNA was performed as described by Sambrook *et al.* (1989). The DNA was resuspended after purification in 20 µl of autoclaved Milli-Q purified water or TE (10:0.1; Section 2.4.1). This method produced DNA of higher quality and was used when required by downstream applications.

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

This method was performed as described in the protocol supplied with the *Taq* DyeDeoxy Terminator Sequencing Kit (Perkin Elmer). The DNA produced by this procedure was used for automated sequencing.

2.5.4 Qiagen High Quality Column Purification of Plasmid DNA (mini prep).

This method was performed to obtain up to 20 µg of high quality DNA to be used for automated sequencing as described in the kit protocol (Qiagen), as modified from the

protocol of Sambrook *et al* (1989). The product was resuspended in 40 µl of autoclaved Milli-Q purified water or TE (10:0.1; Section 2.4.1).

2.5.5 Purification of Fungal genomic DNA.

This protocol for the extraction of high molecular weight DNA was based on the procedures of T. Al-Samarrai and J. Schmid (Al-Samarrai 1999). The preparation is simple, rapid, and produces pure, easily digestible DNA.

Firstly 30 mg of freeze dried mycelia were ground to a fine powder with liquid nitrogen in a 1.5 ml microcentrifuge tube using a microcentrifuge tube grinder. The powder was resuspended in 500 µl of freshly prepared lysis buffer (Tris-acetate (pH 7.8), 40 mM; Na_2EDTA, 20 mM; SDS (w/v) 1%). This was mixed by VIGOROUS pipetting. 164 µl of 5 M NaCl was added to precipitate cellular debris, mixed by inversion and centrifuged at 4 °C (15 000 g in a microcentrifuge) for 20 minutes. The DNA was extracted with one volume of chloroform, centrifuged at 15 000 g for six minutes, then precipitated in 2 volumes of ice cold 95% ethanol and centrifuged at 15 000 g for 5 minutes. The pellet was washed three times with 500 µl of 70% ethanol, dried and resuspended in 50 µl TE (10:0.1; Section 2.4.1).

2.5.6 Lambda Phage DNA preparation.

This method is a modification of the Liquid Lysate method of phage preparation (Sambrook *et al.* 1989) which was first described by Leder *et al.* (1967). Firstly the phage was titred. Then 100 µl of phage particles diluted to a concentration of 10^6-10^7

phage/100 µl were added to 100 µl of plating cells and allowed to adsorb for 30 minutes at 37 °C. This mixture was then transferred to 50 ml of NZCYM (Section 2.2.4) in a 500 ml flask and shaken vigorously at 37 °C for 6-8 hours until lysis occurred. Upon clearing a few drops of chloroform were added and the culture was shaken for a further 15 minutes to lyse any remaining cells. The culture was then transferred to sterile 250 ml Nalgene centrifuge tubes and centrifuged for 10 minutes at 16 300 g (in Sorval GSA rotor equates to10 000 r.p.m.) at 4 °C. The lysate was transferred to a fresh tube where RNAse (10 mg/ml, Section 2.4.7) and DNase (10 mg/ml, Sigma) were added to a final concentration of 10 µg/ml and the preparation was incubated for 1 hour at 37 °C. NaCl and PEG 6000 were added to give final concentrations of 0.5 M and 10% (w/v), and the preparation was put on ice for 2 hours to precipitate the phage. The phage were pelleted by centrifugation at 4 920 g (5 500 r.p.m, GSA), dried, and resuspended in 1 ml of SM buffer (Section 2.4.13), transferred to an eppendorf, then microcentrifuged at 13 000 r.p.m. for 10 minutes. The supernatant was treated with proteinase K (Sigma) to a final concentration of 0.1 mg/ml and incubated at 37 °C for 30 minutes. This was then extracted twice with an equal volume of phenol/chloroform, vortexed for 20 minutes, and centrifuged at 16 300 g for 5 minutes. An equal volume of chloroform was added, vortexed for 10 minutes and centrifuged for 5 minutes. The aqueous phase was then separated and the DNA was precipitated by the addition of 2 volumes of absolute ethanol + 0.3 M ammonium acetate. This was then microcentrifuged for 10 minutes and the pellet washed with 70% ethanol, dried, and resuspended in 200 µl of TE buffer 10:0.1 (Section 2.4.1) and following resuspension debris was removed by a 10 minute microcentrifugation at 13 000 r.p.m.

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 (15 000 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 and DNA was isolated by ethanol precipitation (Section 2.6.2).

2.6.2 Ethanol Precipitation.

In order to concentrate DNA solutions the DNA was precipitated by the addition of 0.1 volume of 3 M sodium acetate and either 2.5 volumes of ethanol or 0.6 volumes of isopropanol followed by a 15 minute incubation on ice. The DNA was pelleted by centrifugation at 15 000 g for 10 minutes and washed with 70% ethanol. The pellet was then dried under vacuum and resuspended in either Milli-Q purified water or TE buffer 10:0.1 (Section 2.4.1).

2.6.3 Agarose Gel Purification.

The DNA to be purified was run on a low melting point agarose gel in TAE buffer (Sections 2.4.2 and 2.9). The gel was then stained in an ethidium bromide solution (Section 2.4.6) and the band of interest excised under long wave U.V light using a sterile

scalpel. DNA was extracted from the fragment using the QIAquick Gel Extraction Kit (Qiagen) according to the supplied protocol.

2.7 DETERMINATION OF DNA CONCENTRATION.

2.7.1 Determination of DNA Concentration by Spectrophotometric Assay.

This procedure was used to determine the purity of the sample as well as the concentration. Each sample was diluted appropriately and the absorbance measured at both 260 nm and 280 nm. The concentration was calculated based on the assumption that an absorbance of 1.0 at 260 nm is equivalent to 50 μg/ml. The degree of purity was indicated by the ratio of absorbance 260 nm / 280 nm, where pure DNA has a ratio of 1.8.

2.7.2 Determination of DNA Concentration by Fluorometric Assay.

This was selected as the best method of DNA quantification for samples to be used for automated sequencing as it produced a consistent specific value for DNA concentration and samples quantified by this method produced the highest quality sequence. DNA was quantified on a Hoefer Scientific TKO 100 Fluorometer according to the protocol supplied by the manufacturer. The scale of the fluorometer was set to 100 using 2 μl of 100 μg / ml calf thymus DNA added to 2 ml of dye solution containing 1 x TNE buffer (Section 2.4.6) and 0.1 μg/ml Hoescht 33258 dye. Once the scale was reproducibly set 2 μl of sample was added to 2 ml of dye solution and the value recorded as the concentration of DNA in ng / μl. This was repeated in triplicate for each sample.

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 then estimated by comparing the intensity of the ethidium bromide fluorescence to that of the concentration standards. This method was used in conjunction with the above methods as it confirmed the integrity of the DNA of interest.

2.8 RESTRICTION ENDONUCLEASE DIGESTION OF DNA.

Restriction endonuclease digests were performed at 37 °C in the buffer specified by the manufacturer of the enzyme. Digest times varied depending on the active life of the restriction endonuclease from a minimum of 1 hour and a maximum of overnight. The digestion was performed with 1-10 units of enzyme per μg of DNA and the enzyme volume never exceeded 10% of the total reaction volume (to maintain low glycerol concentrations). When necessary RNA was removed simultaneously by the addition of RNase A (Section 2.4.6) to a final concentration of 1.0 μg/ml. Digests were confirmed by visualisation of an aliquot on an agarose gel (Section 2.9). If digestion was incomplete more enzyme was added and the sample was incubated further. Completed digests were stored at –20 °C.

2.9 AGAROSE GEL ELECTROPHORESIS.

DNA fragments were size fractionated by electrophoresis through 0.75% - 2.5% agarose dissolved in 1 x TAE buffer (Section 2.4.2) or 1 x TBE buffer (section 2.4.3). Horizon 58 gel platforms requiring 20 ml of agarose were run at 65-95 Volts. Large gels, poured on a Horizon 11.14 gel platform were run overnight at 25 Volts. Gel loading buffer (Section 2.4.9) was added to a final concentration of 1x and dyes present allowed for the estimation of DNA migration. After electrophoresis the gel was submerged in ethidium bromide (Section 2.4.6) for 10-30 minutes, depending on the size of the gel, to stain the DNA. The gel was then washed in water and DNA observed under short wave U.V. Light and photographed.

2.10 DETERMINATION OF DNA MOLECULAR WEIGHTS.

The size of DNA fragments was determined by running them on an agarose gel (Section 2.9) alongside a ladder of known size standards. The molecular weight of the unknown was then determined by measuring the distance migrated by the known fragments and comparing the relative mobility to that of the unknown.

2.11 DNA SUBCLONING.

Fragments of DNA to be subcloned were isolated by gel extraction (Section 2.5.3) and ligated (Section 2.11.1) into the appropriate vector and transformed into an appropriate *E.*

coli host (Sections 2.11.2, 2.11.3). The transformed cells were spread onto supplemented LB plates (Section 2.2.1) containing ampicillin, X-gal, and IPTG and screened using blue/white colour selection. The presence of correct insert was always determined by plasmid DNA extraction (Sections 2.5.1 - 2.5.4), restriction endonuclease digestion (Section 2.8), and gel electrophoresis followed by molecular weight determination (Sections 2.9, 2.10).

2.11.1 SAP Treatment of Plasmid.

2-5 μg of plasmid DNA (high quality, Sections 2.5.2-2.5.4) was digested (Section 2.8) and purified (Sections 2.6.1, 2.6.2). 1-2 μl of Shrimp Alkaline Phosphatase (Boehringer, 1 U/μl) was added to the DNA (in 34 μl H_2O), along with 1 x SAP buffer (Boehringer) to produce a reaction volume of 40 μl. This was incubated at 37 °C for 1 hour and the enzyme deactivated by incubation at 65 °C for 15 minutes. This was then purified as the original plasmid.

2.11.2 DNA Ligations.

Ligation reactions were performed in a total reaction volume of 20 μl containing 2 μl of 10 x ligase buffer (New England Biolabs), 0.5 μl of T4 DNA Ligase (New England Biolabs) and a 2:1 molar ratio of insert : vector with a minimum of 20 ng of vector which was SAP treated (Section 2.11.1) unless both vector ends were different restriction sites. Ligation reactions were placed at 4 °C overnight. Ligations were checked by the comparison of before and after ligation aliquots on an agarose gel (Section 2.9)

2.12 TRANSFORMATION OF *E. coli.*

2.12.1 Calcium Chloride Transformation (heat shock).

The preparation of calcium chloride competent cells and their subsequent transformation was carried out according to Cohen *et al.* (1972) and Ausubel *et al.* (1994).

2.12.2 Electroporation.

The preparation of electroporation competent cells and subsequent transformation were carried out as described by Ausubel *et al.* (1994).

2.13 SOUTHERN BLOTTING AND HYBRIDISATION.

2.13.1 Southern Blotting (Capillary).

The DNA to be blotted was electrophoresed through a large 1-1.5% agarose gel, stained in ethidium bromide (Section 2.9) and photographed with a ruler next to the gel. The gel was then gently agitated in 250 mM HCl for 15 minutes to depurinate the DNA so that high molecular weight regions would transfer efficiently. The DNA in the gel was then denatured by gentle agitation in 500 mM NaOH, 500 mM NaCl for 30 minutes. The gel was then neutralised by two washes in 500 mM Tris [pH7.4], 2 M NaCl for 20 minutes each. Before placing on the blotting apparatus the gel was washed in 2 x SSC for 5 minutes.

The blotting apparatus was constructed as described by Ausubel *et al.*(1994). After blotting overnight the apparatus was disassembled and the membrane (Hybond-N, Amersham) was treated with short wave U.V. light for 3 minutes to crosslink the DNA to the membrane. The final step after crosslinking was to wash the excess salt off the membrane by washing in 2 x SSC for 5 minutes. The membrane was stored between two sheets of 3MM paper at 4 °C.

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

Probe DNA (25 ng) was labelled with α-^{32}P (Amersham) using the Ready-To-Go DNA Labelling Kit (Pharmacia) according to the manufacturers instructions except the labelling reaction produced better results if incubated for one hour at 37 °C. Unincorporated nucleotides were removed using the ProbeQuant G-50 Micro Columns (Pharmacia) as described by the manufacturers.

2.13.3 Southern Blot Hybridisation.

The membrane to be probed was prehybridised for 2 hours at 65°C (rotating oven) in a glass hybridisation tube containing approximately 7 ml of hybridisation buffer (Section 2.4.10). After prehybridisation the solution was replaced with 7 ml fresh, pre-heated hybridisation buffer and the denatured probe was carefully added to the solution but not directly to the membrane. Hybridisation was performed in a rotating oven at 65°C overnight.

After hybridisation the membrane was washed for 30 minutes 3 times with wash solution (3 x SSC, 0.2% SDS) in the rotary oven. The membrane was then wrapped in gladwrap and placed in an X-ray cassette with intensifying screens and either slow film (Fuji Medical) or fast film (Kodak Scientific Imaging). The films were exposed overnight and, pending signal, were put down for the appropriate time for good exposure. The films were developed in a dark room either by placing the films in developing solution (Kodak HC 110) for 5 minutes and fixing solution (Kodak Rapid-Fixer Solution) for at least three minutes, then rinsed in water and air dried. Alternatively, an automated developer (100Plus Automatic X-ray Processor, All Pro Imaging) was used.

2.13.4 Stripping Probe From Membranes.

Membranes to be stripped were placed in a container with boiling 0.1% SDS and left shaking for 30 minutes. This was repeated twice before checking the filter for signal with a Geiger counter; any signal above background resulted in further stripping. When the signal was no longer detectable by Geiger counter stripping was confirmed by autoradiography as described in Section 2.13.3.

2.14 DNA SEQUENCING.

All sequencing was performed using the MUSeq facility at Massey University Palmerston North. DNA to be sequenced was prepared by either method 2.5.3 or 2.5.4 as these produced DNA of sufficient quality. The sequencing reactions were performed using the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin

Elmer) with 500 ng of template DNA, 3.2 pM primer and 8 µl of terminator ready reaction mix in a total volume of 20 µl. Sequence was then continued using primer walking (Table 2.2 shows all primers used).

Table 2.2 Sequencing Primers.

Primer.	Size (nt)	Tm* (°C)	Sequence (5' to 3')	Source.
pUC/M13Fwd	22	70	GCC AGG GTT TTC CCA GTC ACG A	Perkin Elmer.
pUC/M13Rev	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.
SL-PKS1	21	64	GAT GGA GGT GAT GCC ATC TTC	This Study.
SL-PKS2	21	64	GGA CAT GTT GGC AAG TCT CTC	This Study.
SL-PKS3	20	58	GCT TGC TAT TAC ATG CTT GC	This Study.
SL-PKS4	19	58	CTG CAT CAT GAG AAG GAC C	This Study.
Ver2fep1	21	64	CCA AAC ATC AGC CAC AGC AAG	This Study.
Ver2fep2	21	64	GAC ATG AAC GCT GTT GGA TGG	This Study.
Ver2fep3	20	60	CCT GTT ACT GCC AAC GTA TC	This Study.
Ver2fep4	20	62	CAG CTC GAC ATT GTC TGC TC	This Study.
Ver2rep1	20	62	GAC AAC GAC CTT TGC TCC TC	This Study.
Ver2rep2	19	60	CAG TTG CGC AGG ACG ATT C	This Study.
Ver2rep3	21	60	CGA TAC AAG AAC TTC AAG TCG	This Study.
Ver2rep4	20	60	CAG ACA AGC GAC ATG GTT TG	This Study.
Ver2ep1F	22	68	CAG CCG TGA CAT CCA TAA TTC C	This Study.
Ver2ep1R	22	68	GCA TGC ATG TAA ACG TCC ATC C	This Study.

3. CHARACTERISATION OF THE λCGV2 CLONE.

3.1 INTRODUCTION.

The aflatoxin biosynthetic pathway is used as a model for the synthesis of dothistromin as the end products have similar functional groups and some common pathway intermediates have been tentatively identified. This was used to initiate the process of tentative identification of potential biosynthetic genes.

In an attempt to isolate homologues of the *A. parasiticus ver-1* gene from a *D. pini* genomic lambda library C. Gillman (Gillman 1996) isolated two clones containing *ver* like genes, λCGV1 and λCGV2. Sequence analysis revealed λCGV1 contained a putative ketoreductase comparable to *Aspergillus* sp. *ver1*which was named *dkr*1 (Monahan 1998). The *dkr*1 gene is believed to perform an equivalent function in the dothistromin biosynthetic pathway that the *ver*1 gene performs in the aflatoxin biosynthetic pathway. The λCGV2 clone was shown to cross hybridise in this region but the degree of similarity between the clones was not yet established.

Firstly restriction digestion and secondly Southern analysis were used to determine if the clones were indeed different. Restriction analysis indicated that no two restriction fragments present were identical. However Southern analysis indicated that there was a region of sequence similarity as the 0.8kb *Sal*I fragment of λCGV1 hybridised to the 2.45kb *Bam* HI λCGV2 fragment (Gillman 1996). There are three possible reasons for

this: λCGV2 may contain a similar gene in a different pathway, a similar gene in the same pathway or a region of duplication. Gene duplication is not uncommon in fungal genomes as gene repeats occur frequently and are considered to be an important evolutionary tool. There is a well characterised example of gene duplication in the *A. parasiticus* aflatoxin biosynthetic pathway, and as the probe used to identify the *ver* gene was derived from this species, the advent of a duplicated *ver* gene was quite likely.

The primary goal of my project was to characterise the region of similarity present in the clone λCGV2. The first step in this characterisation was restriction analysis and the construction of a restriction map. This was followed by the subcloning of fragments of interest, which were then sequenced by primer walking. These were assembled into a contig using direct sequencing of the lambda clone to confirm junctions. Sequence analysis was then used to determine the nature of the region of similarity and allow the determination of its potential function. The final conclusion was that the gene was more similar to the melanin biosynthetic gene *phn*1 than to the *ver* like genes so throughout this chapter the gene shall be referred to as *phn*1.

3.2 Restriction Analysis.

The first stage in the λCGV2 analysis was confirmation that the purified λ clone DNA remained uncontaminated, intact and was consistent with previous analysis (Gillman 1996). These factors were confirmed through restriction analysis.

The enzymes the restriction analysis was to be based upon were *Bam*HI, *Eco*RI, *Sal*I, and *Hin*dIII. These were selected as they cleave off the λvector arms but do not cleave within the arms. λCGV2 was digested using each of the four restriction enzymes individually and all combinations of two (Section 2.8). A preliminary comparison of λCGV1 and λCGV2 revealed no similarity at this level, these results were in agreement with the preliminary results of Gillman (Gillman 1996) so data was not shown.

3.3 Restriction Mapping.

Due to the complexity of the lambda clone Southern blotting was required for the production of a restriction map. The Southern blot contained 1.2 μg of λCGV2 DNA per lane. The digestion profile was as follows: *Bam*HI, *Eco*RI, *Sal*I, *Hin*dIII, *Bam*HI/*Eco*RI, *Bam*HI/*Sal*I, *Bam*HI/*Hin*dIII, *Eco*RI/*Sal*I, *Eco*RI/*Hin*dIII and *Sal*I/*Hin*dIII and the digests were incubated and separated by gel electrophoresis (figure 3.1). The Southern blot of this gel was hybridised separately to various α-^{32}P labelled fragments of the λCGV2 clone. The fragments selected to be used as probes were the 1.0kb and 2.45kb *Bam*HI fragments and the 2.0kb *Bam*HI/*Eco*RI fragment of λCGV2. These were selected because the 2.45kb *Bam*HI and 2.0kb *Bam*HI/*Eco*RI fragments hybridised to the 0.8kb *Sal*I fragment, which contained part of the putative *ver*1 like *D. pini* sequence (Gillman 1996) and the 1.0kb *Bam*HI fragment was believed to be flanking this region. All hybridisations and washes were standard (Section 2.13). The results of this Southern blot (figures 3.2, 3.3, 3.4, Table 3.1) were used to construct the restriction map (See Figure 3.5). Interestingly the 2..0 kb and 2.45 kb probes produced the same hybridisation

Table 3.1 Data From Restriction Mapping Of Clone λCGV2.

Restriction digest	Fragment Size (kb)
*Bam*HI	0.82[c], 2.5[ab], 4.6, 7.2, 10.3
*Eco*RI	4.1, 11.2[abc], ~15
*Sal*I	2.3[ab], 5.6, 7.2[c], 11.2, ~15
*Hin*dIII	2.1, 2.7, 4.4, 5.4, 7.2, ~15[ab]
*Bam*HI/*Eco*RI	1.0[c], 2.0[ab], 3.9, 6.3, 11.2
*Bam*HI/*Sal*I	0.59, 0.70[c], 1.9[ab], 4.2, 5.8
*Bam*HI/*Hin*dIII	0.73[c], 0.89, 2.0, 2.3, 2.5[ab], 3.7, 5.0, 5.4
*Eco*RI/*Sal*I	2.1[ab], 3.7, 6.9[c]
*Eco*RI/*Hin*dIII	3.2[abc] *
*Sal*I/*Hin*dIII	0.59, 0.70, 0.89, 2.0, 2.1[ab], 2.6, 2.9, 4.9, 5.4[c]

[a] Indicates fragments hybridising to [α-^{32}P]-labelled 2.0 kb *Bam*HI/*Eco*RI fragment.

[b] Indicates fragments hybridising to [α-^{32}P]-labelled 2.45 kb *Bam*HI fragment.

[c] Indicates fragments hybridising to [α-^{32}P]-labelled 1.0 kb *Bam*HI fragment.

* Indicates incomplete restriction digestion.

Figure 3.1 λCGV2 Restriction Digestion Profile.

Each lane contained 1.2 μg of λCGV2 DNA digested by the restriction enzyme listed next to the lane number as follows.

Lane.	**Restriction Enzyme.**
1	*Bam*HI
2	*Eco*RI
3	*Sal*I
4	*Hin*dIII
5	*Bam*HI/*Eco*RI
6	*Bam*HI/*Sal*I
7	*Bam*HI/*Hin*dIII
8	*Eco*RI/*Sal*I
9	*Eco*RI/*Hin*dIII
10	lambda *Eco*RI/*Hin*dIII

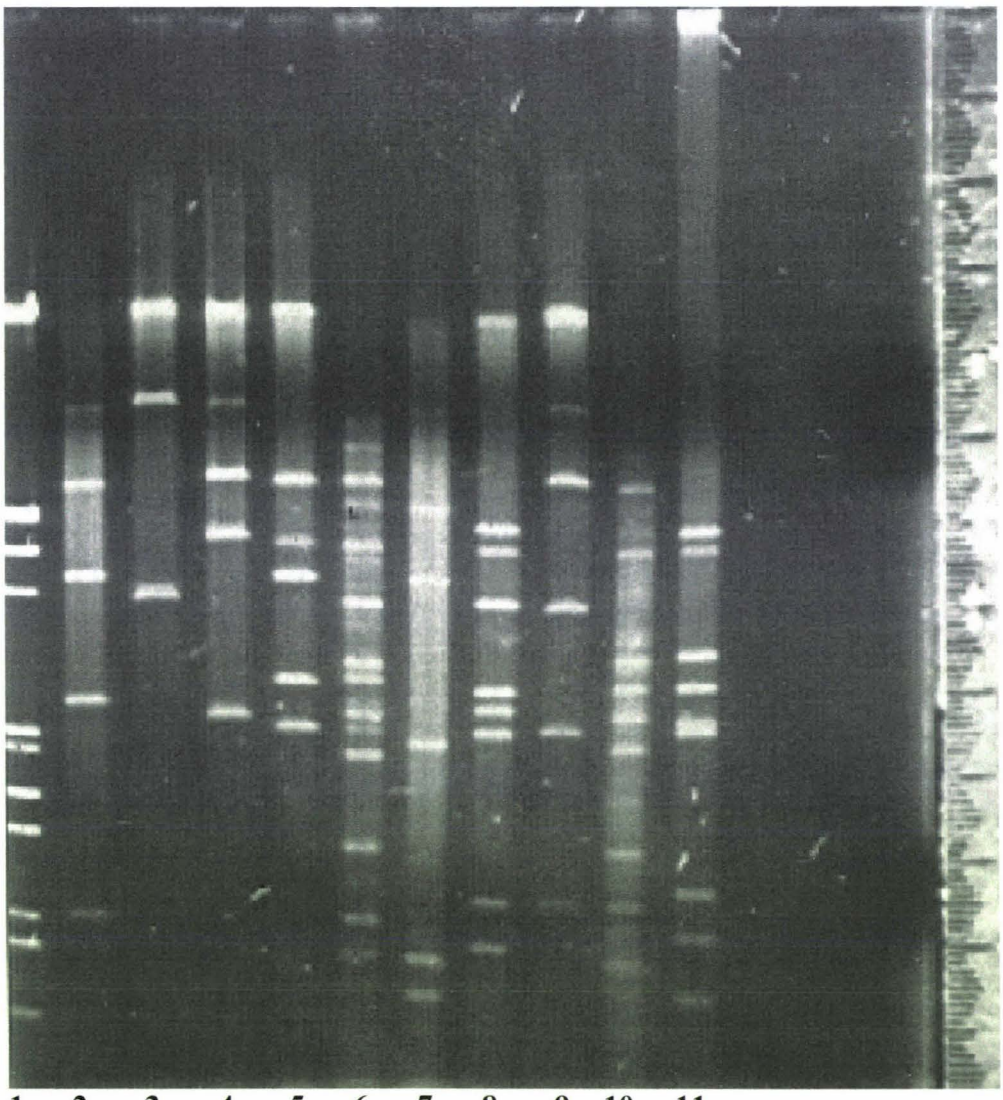

1 2 3 4 5 6 7 8 9 10 11

Figure 3.2 Southern Blot of λCGV2 DNA (Figure 3.1) Probed With 1.0 kb *Bam*HI λCGV2 Fragment.

Lane.	Restriction Enzyme.	Hybridised Band.
1	*Bam*HI	0.8 kb
2	*Eco*RI	11.2 kb
3	*Sal*I	7.2 kb
4	*Hin*dIII	15 kb
5	*Bam*HI/*Eco*RI	1.0 kb
6	*Bam*HI/*Sal*I	0.7 kb
7	*Bam*HI/*Hin*dIII	0.7 kb
8	*Eco*RI/*Sal*I	6.9 kb
9	*Eco*RI/*Hin*dIII	3.4 kb
10	*Sal*I/*Hin*dIII	0.6 kb

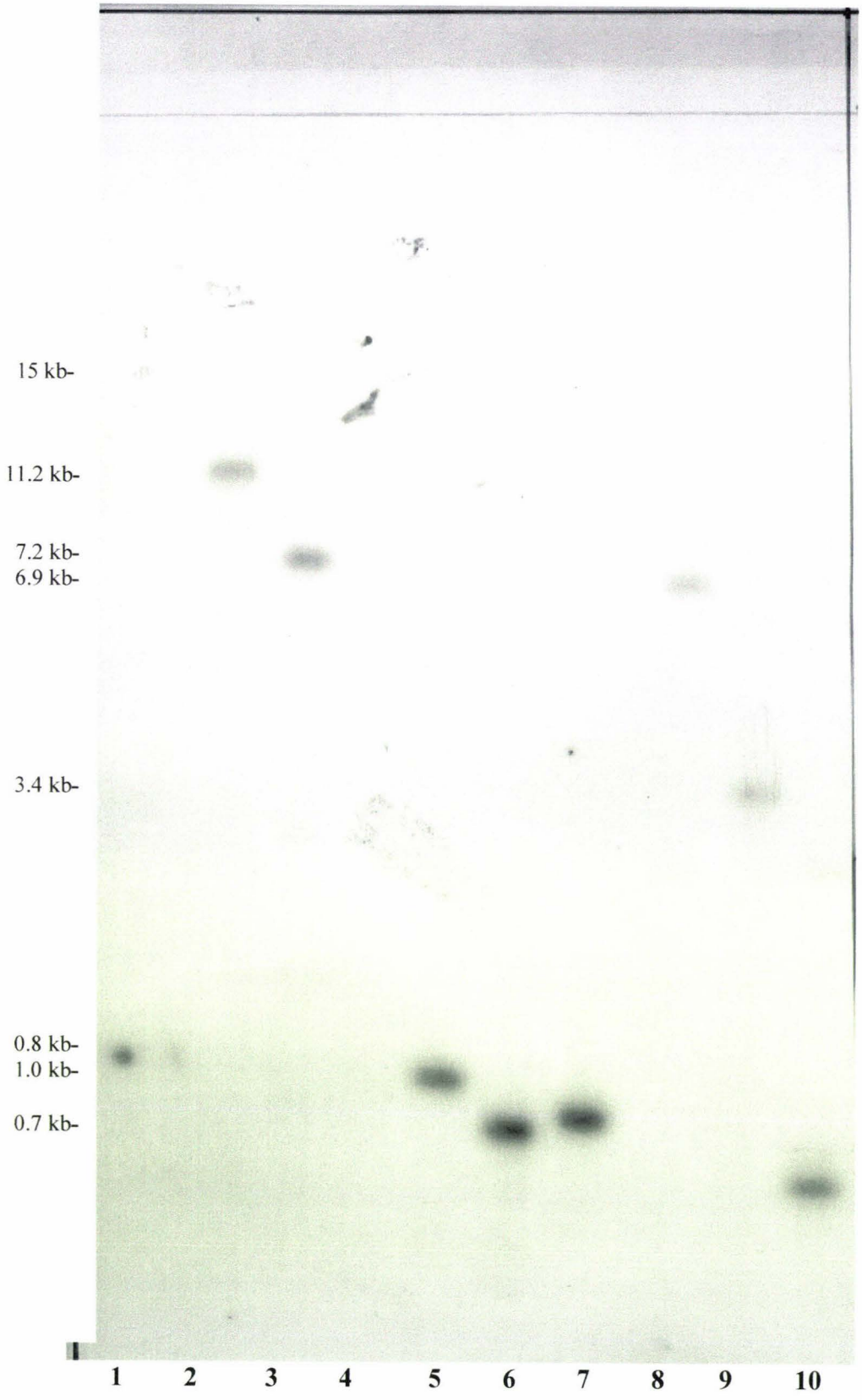

15 kb-
11.2 kb-
7.2 kb-
6.9 kb-
3.4 kb-
0.8 kb-
1.0 kb-
0.7 kb-
1
2
3
4
5
6
7
8
9
10

Figure 3.3 Southern Blot of λCGV2 DNA (Figure 3.1) Probed With 2.45 kb *Bam*HI λCGV2 Fragment.

Lane.	Restriction Enzyme.	Hybridised Band.
1	*Bam*HI	2.5 kb
2	*Eco*RI	11.2 kb
3	*Sal*I	2.3 kb
4	*Hin*dIII	15 kb
5	*Bam*HI/*Eco*RI	2.0 kb
6	*Bam*HI/*Sal*I	1.9 kb
7	*Bam*HI/*Hin*dIII	2.5 kb
8	*Eco*RI/*Sal*I	2.0 kb
9	*Eco*RI/*Hin*dIII	3.2 kb
10	*Sal*I/*Hin*dIII	2.1 kb

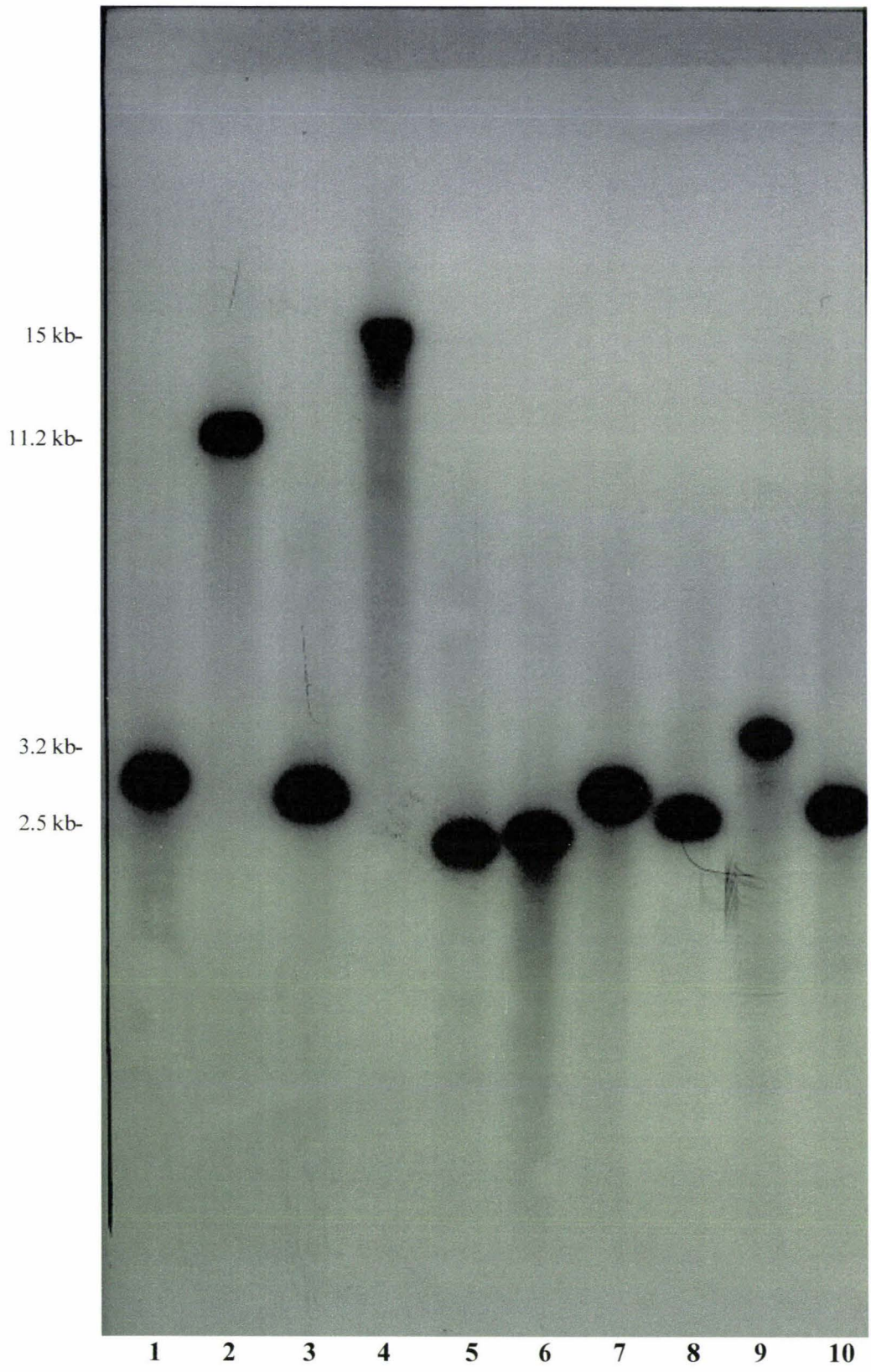

15 kb-
11.2 kb-
3.2 kb-
2.5 kb-
1
2
3
4
5
6
7
8
9
10

Figure 3.4 Southern Blot of λCGV2 DNA (Figure 3.1) Probed With 2.0 kb *Bam*HI/*Eco*RI λCGV2 Fragment.

Lane.	Restriction Enzyme.	Hybridised Band.
1	*Bam*HI	2.5 kb
2	*Eco*RI	11.2 kb
3	*Sal*I	2.3 kb
4	*Hin*dIII	15 kb
5	*Bam*HI/*Eco*RI	2.0 kb
6	*Bam*HI/*Sal*I	1.9 kb
7	*Bam*HI/*Hin*dIII	2.5 kb
8	*Eco*RI/*Sal*I	2.0 kb
9	*Eco*RI/*Hin*dIII	3.2 kb
10	*Sal*I/*Hin*dIII	2.1 kb

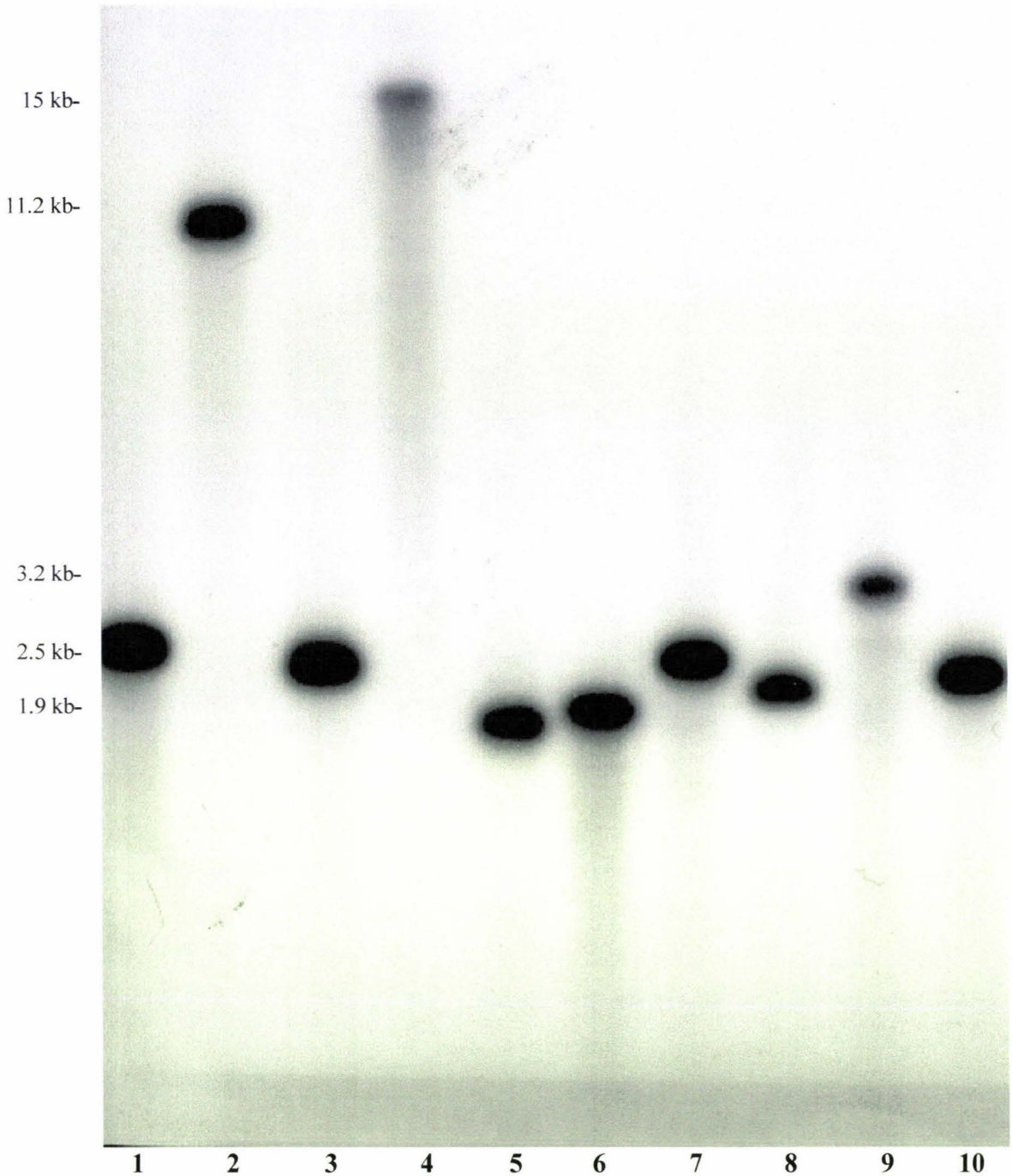
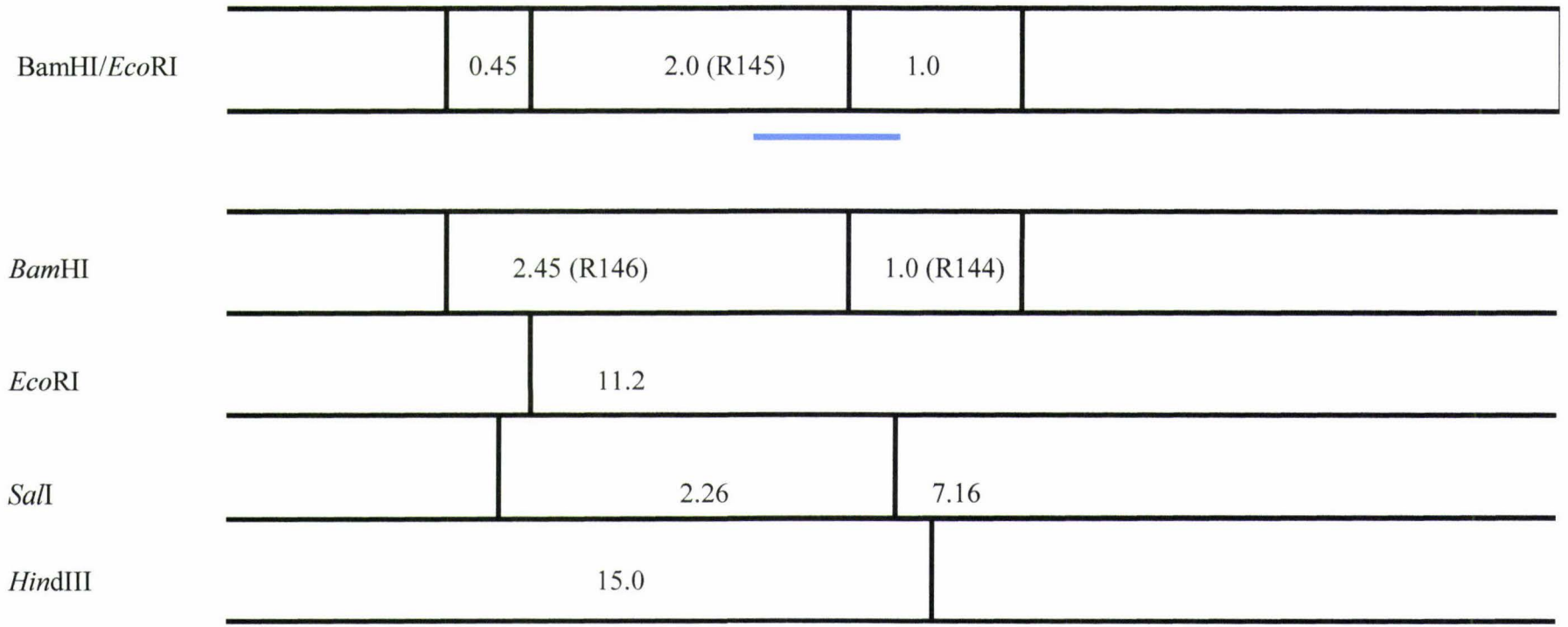

15 kb-
11.2 kb-
3.2 kb-
2.5 kb-
1.9 kb-
1
2
3
4
5
6
7
8
9
10

Figure 3.5 Central Restriction Map of λCGV2.

The λCGV2 restriction map shown was based on Southern hybridisation results and restriction sites confirmed to be present within sequenced regions. All fragment sizes are shown in kb and all known fragments are labelled.

R-numbers indicate subcloned plasmid reference number. (R147 is the 4.4 *Bam*HI fragment and was not included as its position has not been confirmed by sequencing).

The blue line represents the region occupied by the *D. pini phn*1 gene.


BamHI/*Eco*RI
0.45
2.0 (R145)
1.0
*Bam*HI
2.45 (R146)
1.0 (R144)
*Eco*RI
11.2
*Sal*I
2.26
7.16
*Hin*dIII
15.0

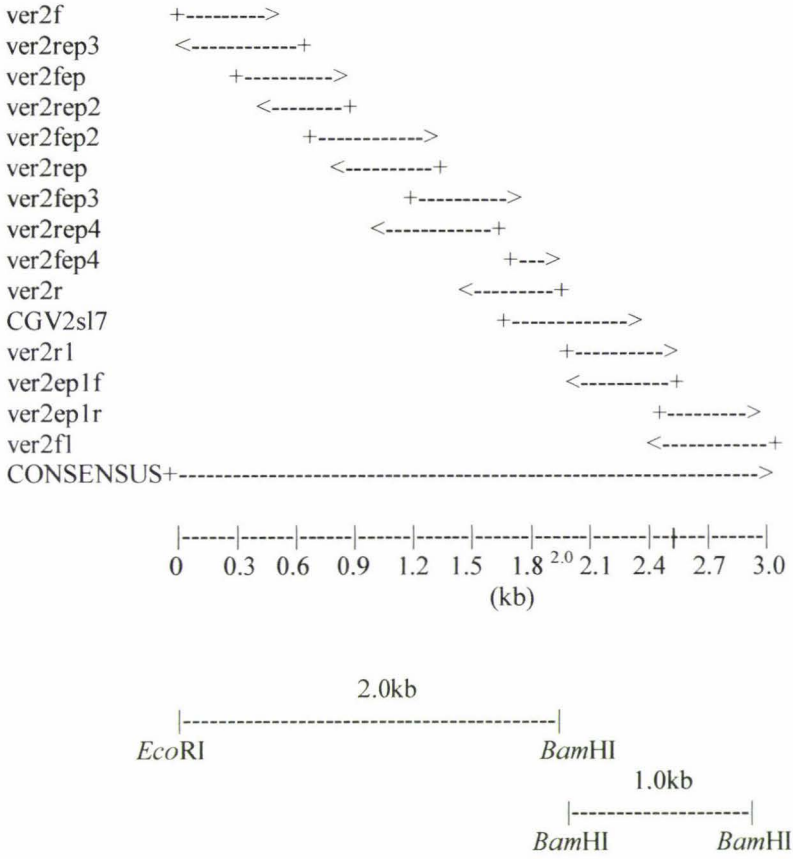
pattern (figures 3.3 and 3.4). Some smaller fragments may have been lost for example an expected 0.45 *Bam*HI/*Eco*RI fragment should have hybridised the 2.45 kb probe and is not apparent. As the 1.0 kb probe was the second applied it is highly unlikely that there was an error and the same blot was exposed twice. The restriction map of the region of interest (region hybridising to 0.8kb *Sal*I fragment of *ver*-1) allowed us to determine a sequencing strategy. Sequence analysis was used to confirm the locations of restriction map sites. The junction between the 2.45 kb *Bam*HI fragment and the 1.0 kb *Bam*HI fragment was sequenced using the primer CGV2sl7 (figure 3.6) and the position of the 4.4 kb *Bam*HI fragment was determined to be at the opposite end of the clone as no junction could be sequenced to the 1.0 kb *Bam*HI fragment.

3.4 Sequencing.

Before sequencing could commence, the fragments of λCGV2 to be sequenced were subcloned (Section 2.11). The subcloned fragments were the *Bam*HI 1.0kb, 2.45kb and 4.4kb fragments and the 2.0 kb *Bam*HI/*Eco*RI fragment (Figure 3.1). These fragments were purified by agarose gel extraction and ligated into pUC 118 which was transformed into *E. coli* XL-1 (Section 2.12.1). The sequencing strategy was to determine preliminary sequence of all three subclones by sequencing (Section 2.14) using the pUC/M13 forward and reverse primers. Once this sequence was complete it was compared to all database sequences at the amino acid level (National Centre for Biological Information, BLAST-X) and any interesting sequences were extended using custom designed primers. (See figure 3.6). This approach yielded the complete sequences of the 1.0kb *Bam*HI and 2.0kb

Figure 3.6 Sequencing Outline for *D. pini phn*I and Surrounding Sequence.

This diagram was obtained from the Big Picture application of the Wisconsin Computing GCG package. The scale shown is in kilobases of DNA sequence and all primer names are listed on the left. CGV2sl7 is the sequence obtained directly from the lambda clone λCGV2 to confirm the junction between the 2.0 kb *Bam*HI/*Eco*RI clone and the 1.0 kb *Bam*HI clone (as shown from left to right) using the ver2fep4 primer.


ver2f
ver2rep3
ver2fep
ver2rep2
ver2fep2
ver2rep
ver2fep3
ver2rep4
ver2fep4
ver2r
CGV2sl7
ver2r1
ver2ep1f
ver2ep1r
ver2f1
CONSENSUS
0 0.3 0.6 0.9 1.2 1.5 1.8 2.0 2.1 2.4 2.7 3.0
(kb)

2.0kb
EcoRI
BamHI
1.0kb
BamHI
BamHI

*Bam*HI/*Eco*RI fragments. This region was then resequenced on the opposite strand to confirm the sequence.

Relevant sequence was then assembled into contigs using the University of Wisconsin Genetics Computer Group (GCG) software programmes. This is how contigs of the sequence obtained from the 1.0kb *Bam*HI and 2.0kb *Bam*HI/*Eco*RI fragments were assembled. The overlap between these two clones was then sequenced directly from the lambda clone to confirm the position and ensure that no sequence was omitted. In this case the primer ver2fep4 was used to generate the sequence CGV2sl7 for this purpose (figure 3.6).

3.5 λCGV2 SEQUENCE ANALYSIS.

3.5.1 Database Comparison.

Analysis, by comparison to known sequence, of the 3kb of the λCGV2 contig using the BLAST database comparison programme revealed the presence of an open reading frame 862bp long (795 bp without intron) encoding a polypeptide of 262 amino acid residues (Figure 3.7). This open reading frame encoded a putative reductase gene. As expected, the BLAST-X results indicated that there was a high degree of similarity to the *A. parasiticus* versicolorin reductase (*ver*-1) gene (74% identity and 83% similarity) to which hybridisation had initially been observed and to the gene equivalents in other

Figure 3.7 *Phn*1 gene sequence and translation.

The numbering on the diagram corresponds to nucleotide number in the λCGV2 sequence listed in Appendix 1.

The Amino acid sequence of the gene is in bold.

The start and termination sequences are labelled in blue.

The intron consensus sequences are labelled in red.

```
      AACTCCCTCCAAGCACACTTCCTCCACCCAACACTACCAACAACAATGTCTGCCTCCCAA
1382   N  S  L  Q  A  H  F  L  H  P  T  L  P  T  T  M  S  A  S  Q

      GCCGTCGAGACCTCCCGCCTGGATGGCAAAGTCGCCCTCGTGACTGGCTCCGGCCGTGGT
1442   A  V  E  T  S  R  L  D  G  K  V  A  L  V  T  G  S  G  R  G

      ATCGGTGCCGCAATGGCTACTGAGCTCGCCCGCCGAGGAGCAAAGGTCGTTGTCAACTAC
1502   I  G  A  A  M  A  T  E  L  A  R  R  G  A  K  V  V  V  N  Y

      GCCAACTTCGCCGAGGCCGCCAACCAGATTGTGGAGCAAATCAAGAAGAACGGCGGCGAT
1562   A  N  F  A  E  A  A  N  Q  I  V  E  Q  I  K  K  N  G  G  D

      GCCATCGCTTTAAAGGCTGATGTCGGCGACGTTGCCCAGACCACCAAGCTCATGGACCAG
1622   A  I  A  L  K  A  D  V  G  D  V  A  Q  T  T  K  L  M  D  Q

      GCCGTTGAGCACTACGGCCAGCTCGACATTGTCTGCTCCAACTCCGGCGTTGTCTCTTTC
1682   A  V  E  H  Y  G  Q  L  D  I  V  C  S  N  S  G  V  V  S  F

      GGCCACTTGAAGGACGTCACTGAGGAGGAGTTCGACCGTGTGTTCCGCATCAACACCCGC
1742   G  H  L  K  D  V  T  E  E  E  F  D  R  V  F  R  I  N  T  R

      GGCCAGTTCTTCGTTGCCCGTGAGGCGTACAAGCACTTGAGCGTTGGTGGTCGCATTATT
1802   G  Q  F  F  V  A  R  E  A  Y  K  H  L  S  V  G  G  R  I  I

      CTTATGGGTTCCATCACTGGTCAGGCCAAGGGTGTGCCAAAGCACGCCGTCTACTCTGGA
1862   L  M  G  S  I  T  G  Q  A  K  G  V  P  K  H  A  V  Y  S  G

      TCCAAGGGCGCCATCGAGACTTTCGTCCGATGCATGGCCATCGACTGTGGTGACAAGAAG
1922   S  K  G  A  I  E  T  F  V  R  C  M  A  I  D  C  G  D  K  K

      ATCACTGTCAACTGCGTTGCTCCAGGTGGTATCAAGACTGACATGTACCACGCTGTATGC
1982   I  T  V  N  C  V  A  P  G  G  I  K  T  D  M  Y  H  A  V  C

      CGCGAATACATTCCGAACGGTGAGAACCTCAGCAACGATCAGGTCGACGAGGTAAGCATA
2042   R  E  Y  I  P  N  G  E  N  L  S  N  D  Q  V  D  E  V  S  I

      ATCCACGCTGGTGGTCATGATATACACATTGCTAATGAGCCGTCGATAGTACGCCAAGAC
2102   I  H  A  G  G  H  D  I  H  I  A  N  E  P  S  I  V  R  Q  D

      GTGGTCTCCAATGCAGCGTGTTGGCCAGCCAATCGACATTGCCCGTGTTGTCTGCTTCCT
2162   V  V  S  N  A  A  C  W  P  A  N  R  H  C  P  C  C  L  L  P

      CGCCTCCCAAGACGGCGAATGGGTCAACGGCAAGGTCATTGGCATTGATGGTGCTGCATG
2222   R  L  P  R  R  R  M  G  Q  R  Q  G  H  W  H  *  W  C  C  M

      CATGTAAACGTCCAGTCCTCGTGGCTTGGAACGACTTTTGGCTGGCAGCGTGTTCACGAG
2282   H  V  N  V  Q  S  S  W  L  G  T  T  F  G  W  Q  R  V  H  E
```

Aspergillus sp.(including 66% identity and 78% similarity to the *A. nidulans stc*U) This supported the assumption that the gene was involved in the dothistromin biosynthetic pathway due to its similarity to the aflatoxin biosynthetic pathway. However the highest degree of similarity apparent in the BLAST-X results was to the *Cochliobolus heterostrophus* polyhydroxynapthalene (*phn*) reductase (77% identity and 86% similarity) and other *phn* homologues in *M. grisea* (75% identity and 86% similarity) and *C. lagenarium* (75% identity and 85% similarity). Hence the *D. pini* ketoreductase may encode an enzyme required for melanin biosynthesis rather than being required for dothistromin biosynthesis, however this is also of relevance to *D. pini's* role as a pathogen as melanin production has been shown to be a key pathogenicity factor in some fungal pathogens being required to establish turgor pressure in the penetration peg of fungal appressoria. The λCGV2 ketoreductase gene was named *phn*1 as this was the closest match presented by the database, yet its function remains uncertain. Although this may also be of relevance to pathogenicity as it was not relevant to the current project sequencing of this clone was suspended.

3.5.2 Sequence Comparisons.

The degree of similarity present in the database sequence prompted further amino acid sequence comparison. The application CLUSTAL W (1.74) Multiple Sequence Pile-up was applied for sequence alignment. Sequences for comparison were selected due to

their similarity to the *D. pini phn*1 gene and grouped into *ver* and *phn* (melanin) reductase like categories (figure 3.7). The sequence alignment scores favour the melanin

biosynthetic genes as most similar with *C. heterostrophus* (aligned score 75), *M. grisea* (aligned score 74) and *C. lagenarium* (aligned score 70) the three highest scoring comparisons all being confirmed melanin biosynthetic genes (See Table 3.2).

3.5.3 GC Content.

The GC content of the coding sequence of the *phn*1 gene was determined using the University of Wisconsin Genetics Computer Group (GCG) composition programme. This calculated a GC content of 57.66 %. The intron sequence only contained 46% GC, this is expected to be lower as there is no selection pressure at the codon level in these non-coding regions. These observations are further elucidated in the discussion section of this chapter.

3.5.4 Codon Usage.

The results for the codon usage in the *phn*1 coding region are presented in appendix 2. These results indicate a bias towards G in the third nucleotide position which was present in 11 of the 24 (46%) most utilised codons. 51 of the 61 sense codons were used, this pattern was also observed for is the codon usage in the *D. pini dkr*1 gene (Monahan 1998).

Figure 3.8 Alignment Of The Deduced Amino Acid Sequence Of *phn*1 With Other *ver* And *phn* Type Genes.

The deduced amino acid sequence of the *D. pini phn*1 gene (dpinPhn1) was aligned with that of the three best melanin type gene matches: *Magnaporthe grisea thn*R (Mgrisea), *Colletotrichum lagenarium thr*1 (Clagenar), *Cochliobolus heterostrophus brn*1 (CHbrn1) and three best *ver* matches *Emericella nidulans stc*U (EnidSTCU), *Aspergillus parasiticus ver* 1 (Apver-1) and the *Dothistroma pini dkr*1 (DpinDkr1).

Sequence alignment was produced using the CLUSTAL W (1.74) Multiple Sequence Alignment Pileup software.

The * symbol indicates all residues in a column are identical.

The melanin gene sequences are in red and *ver* genes are in blue. The *D. pini phn*1 gene is highlighted in green where the sequence is either common to all sequences compared or no common sequence is present, however in areas of *phn* or *ver* conserved sequence it is colored red or blue depending on which conserved residue is present. The residues which are black and bolded immediately precede an intron sequence (note: the *phn*1 sequence was not highlighted at the intron position).

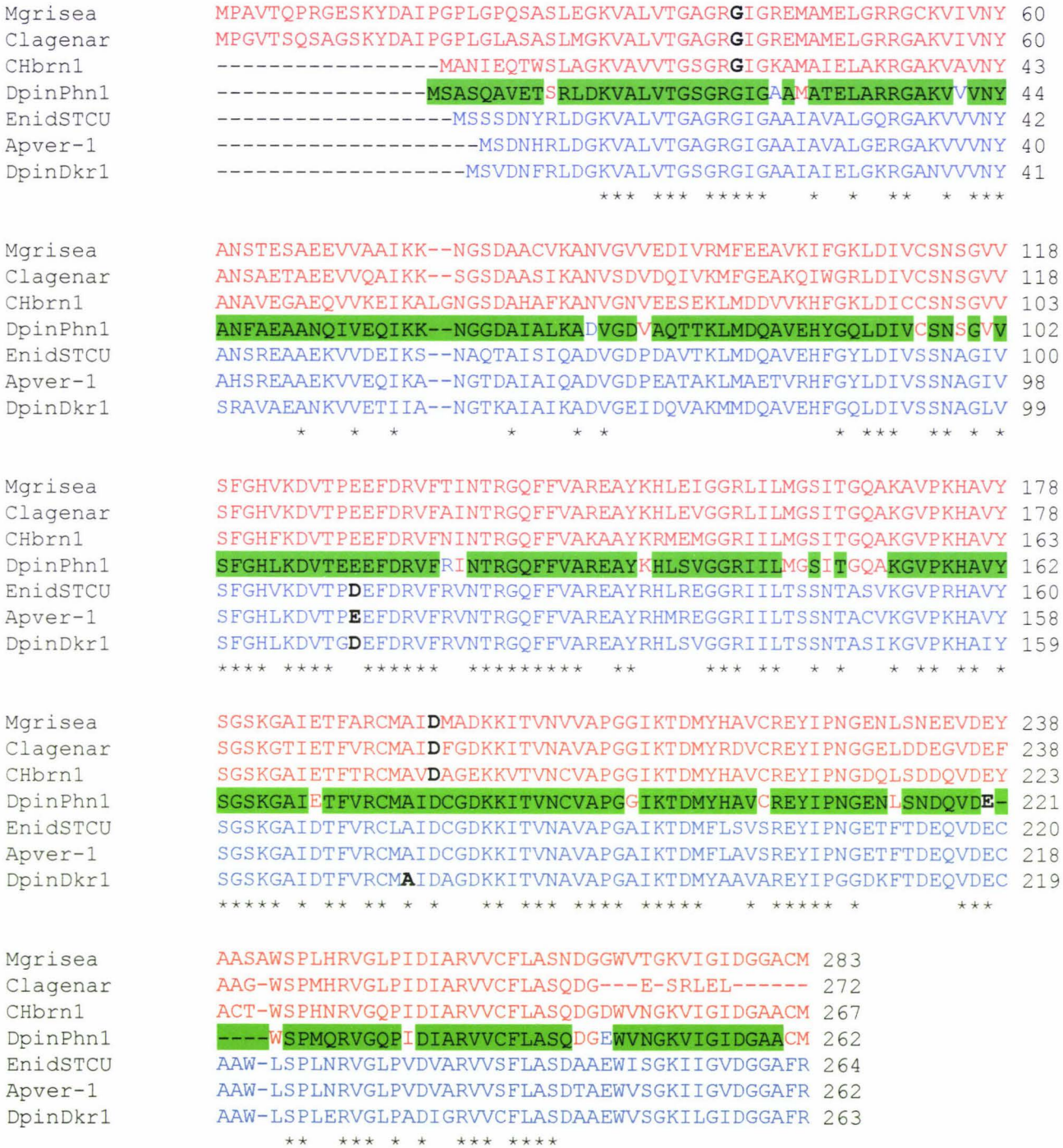

Mgrisea MPAVTQPRGESKYDAIPGPLGPQSASLEGKVALVTGAGRGIGREMAMELGRRGCKVIVNY 60
Clagenar MPGVTSQSAGSKYDAIPGPLGLASASLMGKVALVTGAGRGIGREMAMELGRRGAKVIVNY 60
CHbrn1 ----------MANIEQTWSLAGKVAVVTGSGRGIGKAMAIELAKRGAKVAVNY 43
DpinPhn1 ----------MSASQAVETSRLDKVALVTGSGRGIGAAMATELARRGAKVVVNY 44
EnidSTCU ----------MSSSDNYRLDGKVALVTGAGRGIGAAIAVALGQRGAKVVVNY 42
Apver-1 ----------MSDNHRLDGKVALVTGAGRGIGAAIAVALGERGAKVVVNY 40
DpinDkr1 ----------MSVDNFRLDGKVALVTGSGRGIGAAIAIELGKRGANVVVNY 41
ANSTESAEEVVAAIKK--NGSDAACVKANVGVVEDIVRMFEEAVKIFGKLDIVCSNSGVV 118
ANSAETAEEVVQAIKK--SGSDAASIKANVSDVDQIVKMFGEAKQIWGRLDIVCSNSGVV 118
ANAVEGAEQVVKEIKALGNGSDAHAFKANVGNVEESEKLMDDVVKHFGKLDICCSNSGVV 103
ANFAEAANQIVEQIKK--NGGDAIALKADVGDVAQTTKLMDQAVEHYGQLDIVCSNSGVV 102
ANSREAAEKVVDEIKS--NAQTAISIQADVGDPDAVTKLMDQAVEHFGYLDIVSSNAGIV 100
AHSREAAEKVVEQIKA--NGTDAIAIQADVGDPEATAKLMAETVRHFGYLDIVSSNAGIV 98
SRAVAEANKVVETIIA--NGTKAIAIKADVGEIDQVAKMMDQAVEHFGQLDIVSSNAGLV 99
SFGHVKDVTPEEFDRVFTINTRGQFFVAREAYKHLEIGGRLILMGSITGQAKAVPKHAVY 178
SFGHVKDVTPEEFDRVFAINTRGQFFVAREAYKHLEVGGRLILMGSITGQAKGVPKHAVY 178
SFGHFKDVTPEEFDRVFNINTRGQFFVAKAAYKRMEMGGRIILMGSITGQAKGVPKHAVY 163
SFGHLKDVTEEEFDRVFRINTRGQFFVAREAYKHLSVGGRIILMGSITGQAKGVPKHAVY 162
SFGHVKDVTPDEFDRVFRVNTRGQFFVAREAYRHLREGGRIILTSSNTASVKGVPRHAVY 160
SFGHLKDVTPEEFDRVFRVNTRGQFFVAREAYRHMREGGRIILTSSNTACVKGVPKHAVY 158
SFGHLKDVTGDEFDRVFRVNTRGQFFVAREAYRHLSVGGRIILTSSNTASIKGVPKHAIY 159
SGSKGAIETFARCMAIDMADKKITVNVVAPGGIKTDMYHAVCREYIPNGENLSNEEVDEY 238
SGSKGTIETFVRCMAIDFGDKKITVNAVAPGGIKTDMYRDVCREYIPNGGELDDEGVDEF 238
SGSKGAIETFTRCMAVDAGEKKVTVNCVAPGGIKTDMYHAVCREYIPNGDQLSDDQVDEY 223
SGSKGAIETFVRCMAIDCGDKKITVNCVAPGGIKTDMYHAVCREYIPNGENLSNDQVDE- 221
SGSKGAIDTFVRCLAIDCGDKKITVNAVAPGAIKTDMFLSVSREYIPNGETFTDEQVDEC 220
SGSKGAIDTFVRCMAIDCGDKKITVNAVAPGAIKTDMFLAVSREYIPNGETFTDEQVDEC 218
SGSKGAIDTFVRCMAIDAGDKKITVNAVAPGAIKTDMYAAVAREYIPGGDKFTDEQVDEC 219
AASAWSPLHRVGLPIDIARVVCFLASNDGGWVTGKVIGIDGGACM 283
AAG-WSPMHRVGLPIDIARVVCFLASQDG---E-SRLEL------ 272
ACT-WSPHNRVGQPIDIARVVCFLASQDGDWVNGKVIGIDGAACM 267
----WSPMQRVGQPIDIARVVCFLASQDGEWVNGKVIGIDGAACM 262
AAW-LSPLNRVGLPVDVARVVSFLASDAAEWISGKIIGVDGGAFR 264
AAW-LSPLNRVGLPVDVARVVSFLASDTAEWVSGKIIGVDGGAFR 262
AAW-LSPLERVGLPADIGRVVCFLASDAAEWVSGKILGIDGGAFR 263

Table 3.2 Clustal W Alignment Scores.

Organism (gene)	Clustal % Identity to *D. pini phn*1	Accession Number
M. grisea (*phn*R)	74	L223091
C. lagenarium (thr*1)*	*70*	*D83988*
C. heterostrophus (*brn*1)	75	AB001564
A. nidulans (*stc*U)	66	U347401
A. parasiticus (*ver1*)	68	U639941
D. pini (*dkr*1)	68	Monahan1998

3.5.5 Introns.

An intron was detected on the basis of open reading frame anomaly and its location was consistent with the previously identified fungal consensus intron splice site sequences of 5' [GT(A/G/T)NGTY] and 3' [YAG] (Balance 1986). The intron was 67 bp in length and started at position 668 of the gene sequence. The intron position was also confirmed by comparison to the consensus sequence for both the *ver* and *phn* type genes.

3.6 CHAPTER THREE DISCUSSION.

3.6.1 Restriction Analysis.

The restriction digestion of λCGV2 produced fragments the sizes of which were depicted in table 3.1. When those sizes were compared to those generated by digestion of λCGV1 the absence of any common sized bands was noted. This lack of common bands was interpreted as indicating that no large homologous regions were present, hence the two lambda clones were not identical to one another or substantially overlapping. As cross hybridisation to the *ver*1A *A. parasiticus* probe had previously been observed in both clones (Gillman 1996) this indicated that the region of homology present must represent a second region of homology in the genome rather than a second clone containing the same region.

In *A. parasiticus* a large region of duplication led to a second copy of *ver*1A (*ver*1B) which was shown by sequence analysis to contain a nonsense mutation which led to a premature stop codon and truncated polypeptide of no apparent function. This posed the question of whether a similar duplication event had occurred in *D. pini*. However sequence analysis indicated that the *D. pini phn*1 gene does not contain a premature stop codon, and has a different sequence to the *D. pini dkr*1 gene.

The *D.pini ver*1 also demonstrated similarity to other genes including a 62% sequence similarity to the *brn*1 gene of *C. heterostrophus*, a melanin biosynthetic gene. Therefore the region of homology may be a reductase required by a different pathway. To elucidate upon these areas of uncertainty sequence data would be required.

3.6.2 Sequence analysis.

The clone λCGV2 was subcloned, transformed and sequenced. The sequence data was asssembled, and the region which hybridised the *A. parasiticus ver*1 gene was compared to the NCBI database (Section 3.4). The results of this search confirmed the regions hybridisation to the *A. parasiticus ver*1 probe as the entire amino acid sequence was 83% similar (74% identical) to that of the *A. parasiticus ver*1 gene. However the highest degree of similarity was to the melanin biosynthetic *phn* reductase genes of *C. heterostrophus* (80% identity, 89% similarity) and *M. grisea* (75% identity, 86% similarity). This suggested that the *D. pini ver*-like gene within λCGV2 was either a *ver* type (aflatoxin/sterigmatocystin biosynthetic) gene or a *phn* type (melanin biosynthetic)

gene, though a more precise examination of the sequence was required to determine which with any confidence. This comparison was performed in the form of a multiple sequence pileup containing the three most similar genes in each potential class and the *D. pini phn*1 gene (Figure 3.7). The ClustalW scores (Table 3.2) supported the *phn* type genes as most similar to the *D. pini phn*1. Close examination of the pileups showed that overlaps of discrete regions apparently conserved in *ver* type genes or *phn* type genes exist and the *phn*1 gene is grouped with the *phn* type genes (see figure 3.7) in 24 of a total of 29 cases (83%). The sum of these analyses supported the possibility that *phn*1 was a melanin biosynthetic gene. However as melanin biosynthesis had never been reported in *D. pini* an examination of potential roles in dothistromin biosynthesis of this gene was required.

3.6.3 Potential Functions of *D. pini phn*1.

The sequence comparison revealed the closest amino acid match to *D. pini phn*1 to be the *Cochliobolus heterostrophus brn*1 gene. Genetic complementation studies of melanin deficient *C. heterostrophus* mutants indicated that the *brn*1 gene encoded a reductase gene with an amino acid sequence similar to other fungal reductases. Further characterisation showed that *brn*1 was responsible for the reduction of both a 1,3,6,8-tetrahydroxynapthalene and a 1,3,8-trihydroxynapthalene in the melanin biosynthetic pathway (Shimizu 1997). These data are sufficient to propound that the *D. pini phn*1 gene product may also be involved in the melanin biosynthetic pathway.

Based on amino acid sequence similarity to aflatoxin *ver*1 the putative *D. pini ver*1 gene is believed to catalyse the equivalent step in the dothistromin biosynthetic pathway, the conversion of versicolorin A to sterigmatocystin. As *D. pini phn*1 demonstrated a 68% Clustal identity (Table 3.2) to *D. pini dkr*1, the potential for *phn*1 to fill this role requires investigation.

The production of aflatoxin and sterigmatocystin in *Aspergillus* sp. stems from polyketides, as does melanin biosynthesis. Both also include polyphenol dehydroxylation steps for example the versicolorin to sterigmatocystin conversion in aflatoxin. These reactions are all categorised as dehydroxylations of meta-diphenols (Figure 3.9). Extensive studies of the *M. grisea phn* reductase (Vidal-Cross 1994) led to the proposal that the protein coded by *ver*1 in *Aspergillus* sp. was the aflatoxin counterpart of the melanin *phn* reductase.

There is more than one enzyme involved in the versicolorin A to sterigmatocystin conversion. Analysis of the *A. nidulans stc*S (equivalent to *A. parasiticus ver*B) showed that it encoded a cytochrome P-450 monooxygenase and its disruption led to the accumulation of versicolorin A. The disruption of *A. nidulans stc*U (*A. parasiticus ver*1) also led to the accumulation of versicolorin A. This constitutes the first genetic proof that the conversion of versicolorin A to sterigmatocystin requires more than one enzymatic activity. Assuming that the *stc*S activity mimics that of similar P-450s, it is likely that *stc*S catalyses one of the proposed oxidation steps required to convert versicolorin A to

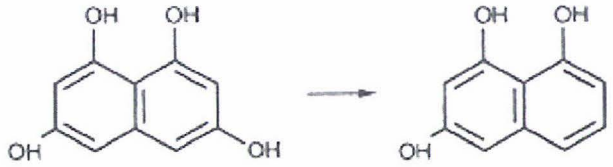
Figure 3.9 Examples of dehydroxylations of meta-diphenols.

T_4HN is tetrahydroxynapthalene.

T_3HN is trihyroxynapthalene.

DHN is 1,8-dihydroxynapthalene.

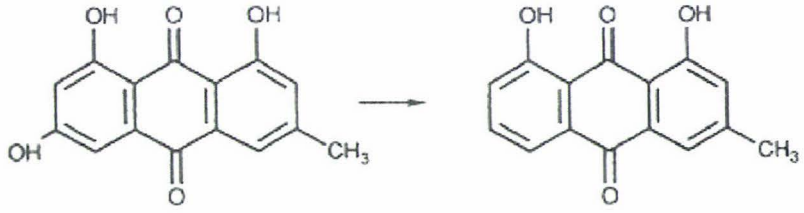
MELANIN



T_4HN T_3HN

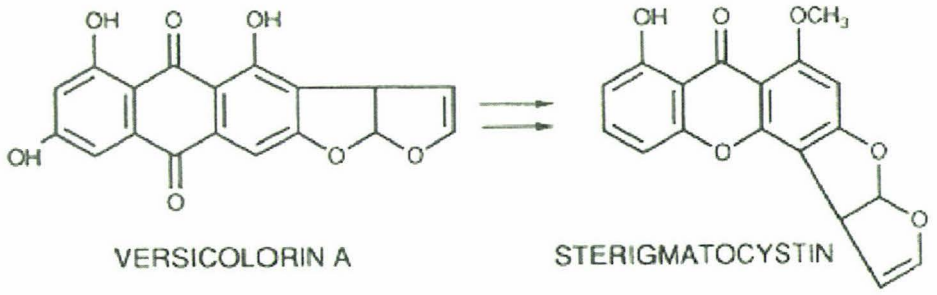
T_3HN DHN

CYNODONTIN



EMODIN CHRYSOPHANOL

AFLATOXIN B_1



VERSICOLORIN A STERIGMATOCYSTIN

sterigmatocystin (Keller 1995). Therefore it is plausible that in *D. pini* more than one enzymatic activity may be responsible for each conversion.

The aflatoxin biosynthetic pathway is used as a model for the synthesis of dothistromin as the end products have similar functional groups and some common pathway intermediates have been tentatively identified. The recent identification of multiple enzymes to catalyse a single conversion (described above, Keller 1995) in the model system leads to the suggestion that perhaps the *D. pini phn*1gene product is a catalyst in the conversion of versicolorin A to sterigmatocystin or to one of the intermediates in the dothistromin biosynthetic pathway as well as being involved in melanin biosynthesis. As the melanin biosynthetic counterpart catalyses two different reductions in the melanin pathway, it is possible that this reductase (*phn*1) also actively partakes in reactions in the dothistromin biosynthetic pathway. This is supported by the multifunctional nature of the melanin reductase genes.

In summary, I propose the tentative hypothesis that *phn*1 as well as *dkr*1 is involved in the same step in dothistromin biosynthesis and that neither are essential as long as the other remains intact. The only way to accurately assess the situation is through transformational knockouts of each and of both followed by accumulation product analysis.

3.6.4 GC Content and Codon Analysis.

The *D. pini phn*1 gene was shown to have a GC content of 57.7%. Codon analysis indicates a bias towards G in the third nucleotide position which was present in 11 of the 24 (46%) most utilised codons. This is consistent with all other *D. pini* genes sequenced to date (table 3.3). The model filamentous fungus, *A. nidulans*, demonstrates a genomic GC content close to 50% indicating little overall mutational bias. However all genes in *A. nidulans* known (or expected) to be highly expressed are highly biased towards a set of 20 optimal codons, most of which are C- or G- ending making highly expressed genes more GC rich than genes which are not highly expressed (Lloyd, 1991). Assuming that a similar trend may exist in *D. pini* we would expect multiple enzymes required for a single conversion to be expressed at similar levels so we would expect a similar GC content due to an equal amount of selection pressure for a G or C at the third codon position.

Previously it was suggested that perhaps *phn*1 and *dkr*1 were both capable of catalysing a single conversion. Continuing this argument we would expect the two to have very similar GC contents, assuming they entered the pathway at a similar evolutionary era. The results indicate that the *dkr*1 gene has a GC content of 56.7% and the *phn*1 gene has a GC content of 57.7%. This is substantially less variation than there is between other putative pathway genes such as *dtp*1 (53.8%) and *dte*1 (57.8%) hence is probably not significant.

Table 3.3: The GC Content of All *D. pini* Genes Sequenced To Date.

Gene Name	Percentage GC Content
*ddh*1	56.6%
*dkr*1	56.7%
*dox*1	54.5%
*dtp*1	53.8%
*dte*1	57.8%
*phn*1-	57.7%
pks[1]	54.5%

[1] For *pks* see section 4.4.3

3.6.5 Intron Analysis.

The position of the first intron of the *D. pini dkr*1 gene is identical to that of the *A. nidulans* and *A. parasiticus ver*1 genes. The melanin biosynthetic *phn*1 genes of *C. lagenarium, M. grisea, and C. heterostrophus* also demonstrate conservation of the first intron position. It follows that we would expect the first intron position of the *D. pini phn*1 gene to be identical to one of these, presumably the *phn*1 gene family. However, our results indicate that the intron position within the *D. pini phn*1 gene is unique (figure 3.8). This is of interest as previous results demonstrate a high degree of conservation to these two types of gene.

3.7 SUMMARY.

The sum of the results of the analysis of λCGV2 demonstrate that the clone is independent of other characterised lambda clones (λCGV1). The gene present in λCGV2 resembling the aflatoxin *ver*1 gene has been shown to demonstrate a greater similarity to the melanin biosynthetic *phn*1 gene family than to the *ver*1 family.

It has been speculated that the *D. pini phn*1 gene may be involved in the dothistromin biosynthetic pathway as well as in melanin biosynthesis but this has not been quantified. The hypothesis requires testing through a process of transformational gene disruptions and intermediate accumulation studies.

The remainder of the λCGV2 clone has not yet been analysed as melanin biosynthesis is not the subject of this project, however further analysis of the flanking regions may shed light on the exact nature of this gene.

4. CHARACTERISATION OF THE POLYKETIDE SYNTHASE CONTAINING CLONE, λBMKSA.

4.1 Introduction.

The *D. pini* genomic library produced by C. Gillman (Gillman 1996) was also screened by B. Morgan (Morgan 1997) in an attempt to isolate further *D. pini* homologues of *Aspergillus* Sp. aflatoxin biosynthetic genes. This second screening was performed using an *Aspergillus parasiticus pks*L1 (polyketide synthase) probe and led to the isolation of the lambda clone λBMKSA.

Partial sequence of the λBMKSA clone was generated and database analysis suggested it contained a putative *pks* domain. However, this sequence was incomplete and contained many anomalies (Morgan 1997). Sequence of the lambda clone λCGV1 had revealed the presence of a thioesterase domain, a potential domain of the *D. pini pks* gene. This led to the proposition that the clones were overlapping or possibly identical. Prior to further analysis of the λBMKSA clone a comparison to the previously characterised lambda clones (λCGV1 and λCGV2) was therefore required to determine if the lambda clones were identical or contained common regions. As this was the first comparison of these three clones it began at the level of restriction digestion (inconclusive) which was followed by a more conclusive comparison through Southern analysis.

The Southern blotting analysis involved the hybridisation of probes derived from each of the lambda clones to restriction digests of each of the lambda clones. Genomic *D. pini* DNA was included to these blots to confirm the presence of the DNA contained in the lambda clones within the genome of *D. pini*. This analysis demonstrated that the lambda clones were not identical and contained sequences which are present in the genome of *D. pini*.

As the lambda clones appeared to be different from one another further analysis of the clone λBMKSA was carried out: specifically the generation of double stranded DNA sequence of the region containing the putative *pks* gene. Primary sequences of subclones of the λBMKSA clone were generated and compared to the Genbank database using the BLAST-X application. Regions demonstrating homology to potential dothistromin biosynthetic genes were sequenced further by primer walking. The results of this sequence analysis were used to determine the presence of putative dothistromin biosynthetic genes.

4.2 Comparison of Lambda clones.

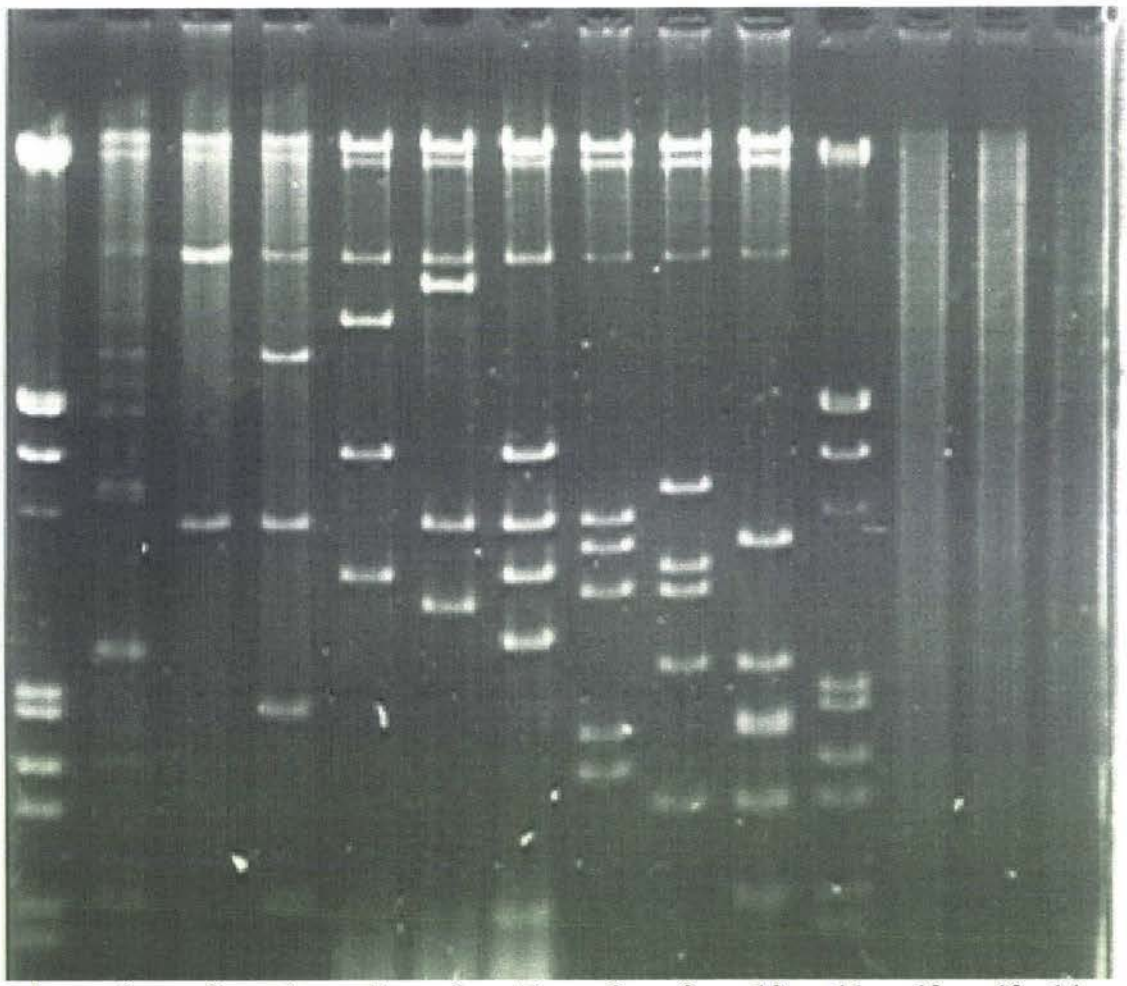
4.2.1 Restriction Digest Comparisons.

Restriction digests of the lambda clones λCGV1, λCGV2 and λBMKSA were compared. Reaction mixtures containing 1.2μg of DNA from each of the clones or 2.5μg of genomic *D. pini* DNA were digested overnight with *Bam*HI, *Eco*RI and *Bam*HI/*Eco*RI (Section

Figure 4.1 Lambda Clone Comparison Restriction Profile.

1.2 μg of each the lambda clones λCGV1, λCGV2 and λBMKSA was digested separately and run along side 3 μg digested *D. pini* genomic DNA. Each of the subjects was digested using *Bam*HI, *Eco*RI and *Bam*HI/*Eco*RI.

Lane.	**DNA Substrate.**	**Restriction Enzyme.**
1	Standard Lambda	*Eco*RI/*Hin*DIII
2	λCGV2	*Bam*HI
3	λCGV2	*Eco*RI
4	λCGV2	*Bam*HI/*Eco*RI
5	λBMKSA	*Bam*HI
6	λBMKSA	*Eco*RI
7	λBMKSA	*Bam*HI/*Eco*RI
8	λCGV1	*Bam*HI
9	λCGV1	*Eco*RI
10	λCGV1	*Bam*HI/*Eco*RI
11	Standard Lambda	*Eco*RI/*Hin*DIII
12	*D. pini* genomic DNA	*Bam*HI
13	*D. pini* genomic DNA	*Eco*RI
14	*D. pini* genomic DNA	*Bam*HI/*Eco*RI



1 2 3 4 5 6 7 8 9 10 11 12 13 14

2.8). These digests were run on an agarose gel (Section 2.9). The restriction digest profile demonstrated the absence of any common insert bands (figure 4.1). This indicated that no clone shares a common *Bam*HI, *Eco*RI or *Bam*HI/*Eco*RI digestion profile with any other hence the clones were not identical and there were unlikely to be long regions of overlap. However regions of similarity beyond the resolution of this protocol, such as overlapping end fragments, may also be present so a more conclusive approach was required.

4.2.2 Southern Blot Comparisons.

Further confirmation that the three lambda clones compared so far: λCGV1, λCGV2 and λBMKSA were independent, representative of the genome (not scrambled or contaminants), and contained no regions of overlap was required. The approach applied to confirm these factors was Southern blotting, and this would also confirm the presence of the lambda clones in the genome of *D. pini*, and potentially reveal whether the clones were contiguous.

The previously described gel (figure 4.1) was blotted overnight and crosslinked to the membrane (Sections 2.13.1). The fixed membrane was then cut to separate the genomic digests from the lambda digests to prevent titration of the probes during hybridisation.

The [α-^{32}P] labelled DNA probes selected were the λCGV1 2.1 kb *Bam*HI/*Eco*RI fragment, λCGV2 2.0 kb *Bam*HI/*Eco*RI fragment and the λBMKSA 2.4 kb

Figure 4.2 Southern Blot of Digested Lambda Clones (Figure 4.1) Probed With Lambda Clone Restriction Fragments.

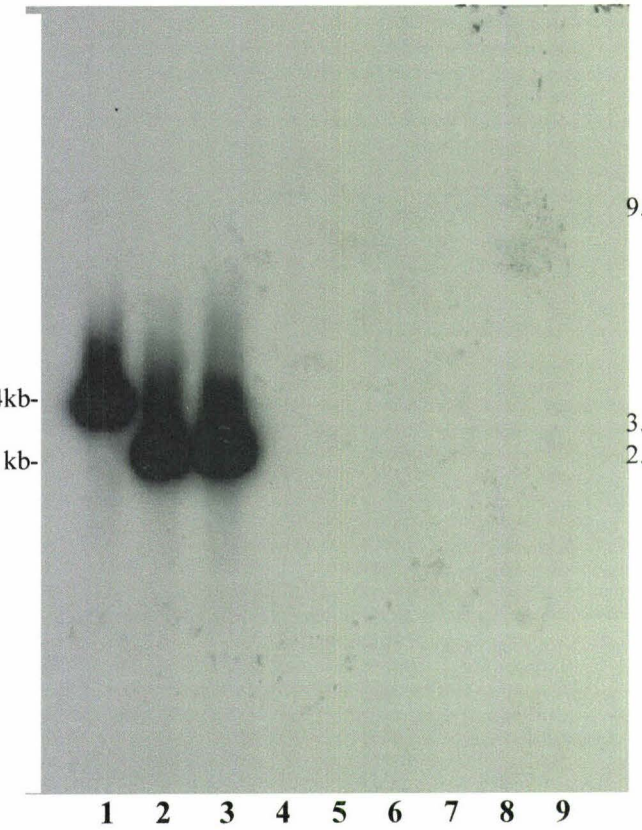
		Probes and Bands Hybridised.		
Lane	**Lambda Digest.**	**λCGV1 2.1 kb BamHI/EcoRI**	**λCGV2 2.0 kb BamHI/EcoRI**	**λBMKSA 2.4 kb BamHI/EcoRI**
1	λCGV1 +B	3.4 kb	-	-
2	λCGV1 +E	2.5 kb	-	-
3	λCGV1 +B/E	2.5 kb	-	-
4	λCGV2 +B	-	2.3 kb	-
5	λCGV2 +E	-	11.2 kb	-
6	λCGV2 B+E	-	2.0 kb	-
7	λBMKSA +B	-	-	9.0
8	λBMKSA +E	-	-	3.0
9	λBMKSA B+E	-	-	2.4

B Represents *Bam*HI

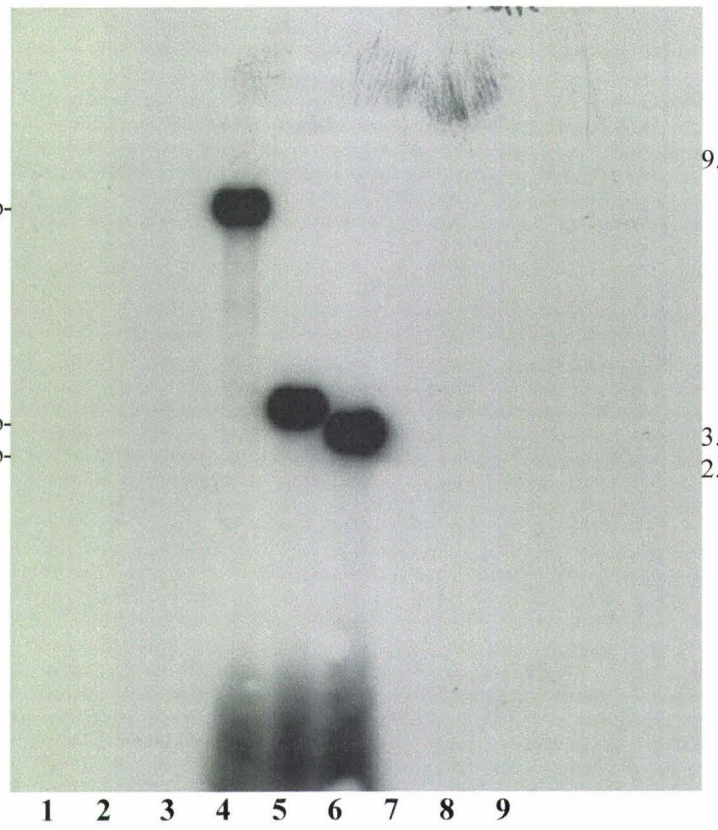
E Represents *Eco*RI

*Bam*HI/*Eco*RI fragment as they are at the very ends of their respective clones (figure 5.8).

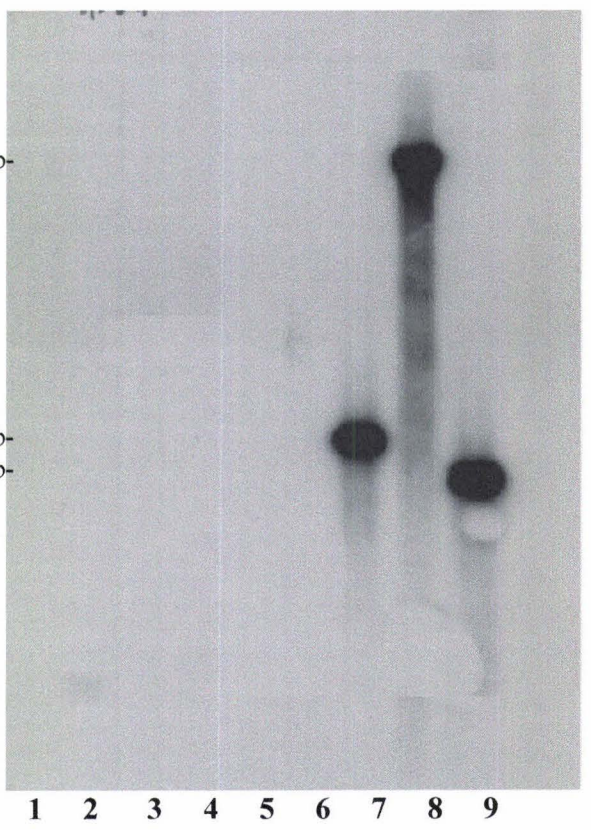
1. λCGV1



2. λBMKSA



3. λCGV2



These probes were prepared (Section 2.13.2) and half was added to the genomic hybridisation and the other half was added to the lambda hybridisation (Section 2.13.3).

All three of the selected probes hybridised only to their respective lambda clones and no cross hybridisation was apparent (figure 4.2). The probes all hybridised to the genomic digests, however the genomic fragments hybridised were of different sizes in each case. The autoradiographs produced for the genomic comparison are not shown as a more thorough comparison between the clones at this level is presented in section 5.

4.3 Sequencing.

Prior to sequencing, the λBMKSA clone required subcloning. The clone was digested using *Bam*HI/*Eco*RI (Section 2.8) to produce fragments of the sizes 0.3kb, 0.94kb, 2.4kb, 3.0kb, 3.5kb and 4.5kb and the fragments were separated in a 1% TAE Agarose gel (Section 2.9). The fragments were gel extracted and purified (Section 2.6.3) before being ligated into the *E. coli* vector pUC118. The 0.3kb fragment was not subcloned as this could easily be sequenced directly from the lambda clone if required. The ligation mixture was then transformed into *E. coli* by electroporation (Section 2.11). Plasmid presence was confirmed by the extraction and digestion of the transformed plasmid from cultured transformed *E. coli* and all plasmids to be sequenced were purified using the Qiagen plasmid miniprep (Section 2.5.4).

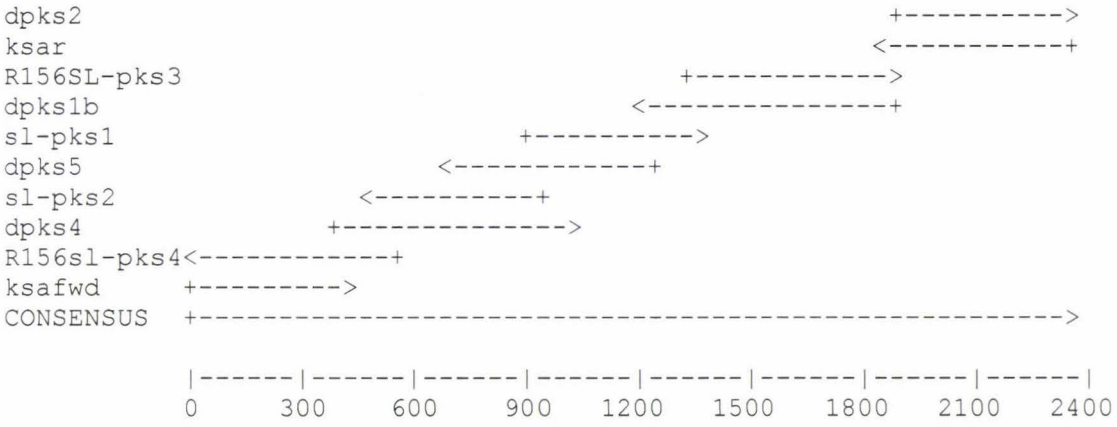
The initial sequence of each of the clones was generated using the standard pUC forward primer, and in the case of the 4.5 kb *Bam*HI subclone (R162) and the 2.4 kb *Bam*HI/*Eco*RI also the pUC reverse primer. These initial sequences (appendix 1) were compared to database sequences at the amino acid level (National Centre for Biological Information, BLAST-X). The 2.4kb *Bam*HI/*Eco*RI fragment contained a region with a high degree of amino acid sequence similarity to the *A. parasiticus pks*1 gene. This is believed to be a key gene in the putative dothistromin biosynthetic pathway so the sequences were extended using custom designed primers. The sequencing outline is depicted in figure 4.3. These two sequences (forward strand and reverse strand) were then assembled into a contig using the University of Wisconsin Genetics Computer Group (GCG) software package.

Another result yielded from this primary sequence was a region in the 3.0 kb *Bam*HI/*Eco*RI fragment of λBMKSA with similarity to the *A. nidulans stc*W gene. This sequence showed a 60% identity and 72% similarity (BLAST-X) to the *stc*W gene present in *A. nidulans*, a putative FAD containing monooxygenase. This adds further support to the hypothesis that the *pks* gene contained in this clone is part of a putative dothistromin biosynthetic cluster.

Preliminary sequences of 4.5 kb *Bam*HI, 0.94 kb *Bam*HI/*Eco*RI and the 3.5 kb *Eco*RI subcloned fragments of the λBMKSA clone did not reveal any further sequences which matched significantly to other genes, however if the gene density is similar to that of the fully characterised λCGV1 clone we would expect further genes to be contained in these

Figure 4.3 Sequencing Outline for *D. pini pks*1.

This diagram was obtained from the Big Picture Application of the Wisconsin Computing GCG package. The scale shown is in kilobases of DNA sequence and all primer names are listed on the left.


dpks2
ksar
R156SL-pks3
dpks1b
sl-pks1
dpks5
sl-pks2
dpks4
R156sl-pks4
ksafwd
CONSENSUS
0
300
600
900
1200
1500
1800
2100
2400

fragments. These partial sequences are present in appendix one. Further primers were designed, however due to a lack of funding the sequence was not generated. Further sequencing of these subcloned fragments would be likely to reveal further potential dothistromin biosynthetic genes.

4.4 *Pks* Sequence analysis.

4.4.1 Database Comparison.

The 2.38kb nucleotide sequence generated from the *Bam*HI/*Eco*RI subcloned fragment of the clone λBMKSA was translated to a predicted amino acid sequence and compared to the National Centre for Biological Information database using the BLAST-X application. The results of this analysis indicated that a 2265 nucleotide open reading frame existed coding for a 755 amino acid sequence (appendix 1). This partial gene sequence demonstrated a 67% identity and 79% similarity to the *A. parasiticus* polyketide synthase (*pks*) and a 64% identity and 75% similarity to the *A. nidulans stc*A gene, also an equivalent polyketide synthase.

4.4.2 Conserved Domains.

Any *pks* gene contains several domains . Those present in the *D. pini* sequence were the β-ketoacyl ACP synthase domain and the acyltransferase domain. These domains were aligned to other equivalent domains in *Aspergillus* sp. and other organisms with greater

sequence divergence as depicted in figure 4.4. This demonstrated that the *D. pini pks* domains shared a high degree of homology with those of other similar fungi, and shared the conserved *pks* domains of more distant species.

4.4.3 GC Content and Codon Usage.

The GC content of the coding sequence of the putative *D. pini pks* gene was determined using the University of Wisconsin Genetics Computer Group (GCG) composition programme. The calculated GC content was 54.53 % (as compared to *phn*1, 57.66 %). This is fairly typical of all *D. pini* genes sequenced to date (table 3.3)

The results for codon usage in the *pks* coding sequence generated so far are presented in appendix 2. These results indicate that 60 of the 61 sense codons were used, however only 8 of the 23 (35%) favoured codons contained a G in the third position, 10 contained a C in the third position (43%). This suggests far less evolutionary selection than the *dkr*1 gene in which only 51 of the sense codons were used with a strong preference for C in the third codon position.

Figure 4.4 Active Site Comparison of *pks* Domains.

The accession numbers for sequences shown refer to the GenBank database. The final sequence of each comparison highlights conserved regions and the bold residue indicates a known active site (Yu 1995).

β-Ketoacyl ACP synthase active sites.

		Accession Number
P. patulum MSAS	GPSTAVDAACASSLVAI	P22367
A. nidulans wA	GPSVSVDTACSSSLAAI	Q03149
A. nidulans pks	GPSYSNDTACSSSLAAI	Q12397
A. parasiticus pks	GPSYTNDTACSSSLAAI	Q12053
D. pini pks	GPSFTNDTACSSSLAAI	This Study
M. tuberculosis bovis MAS	GPAMTFDTACSSGLMAV	CAB06108
S. erythraea ERYA	GPAMTVDTACSSGLTAL	I1702361A
Active site.	GP----DTA**C**SS-L	

(Cystein involved in thioester linkage formation.)

Acyltransferase active sites.

P. patulum MSAS	GITPQAVIGHSVGEIAASVVAGALSPAE
A. nidulans Wa	GITPSFVLGHSLGDFAAMNAAGVLSTSD
A. nidulans pks	GLKPSAVIGHSLGEYAALYISGVLSAAD
A. parasiticus pks	GIRPDVTVGHSLGEFAALYAAGVLSASD
D. pini pks	GIKASAVVGHSLGEYAALYAAGVLSASD
M. tuberculosis bovis MAS	GVRPGAVVGHSMGESAAAVVAGALSLED
S. erythraea ERYA	GVEPAAVVGHSQGEIAAAHVAGALTLED
Active site.	G-------GH**S**-G--A

(Serine involved in acyl-enzyme intermediate formation)

4.5 DISCUSSION.

4.5.2 Sequence Analysis.

4.5.2.1 Putative *Pks* Functions.

The sequence generated from the 2.4kb *Bam*HI/*Eco*RI fragment of the λBMKSA clone demonstrated a 67% identity and 79% similarity to the *A. parasiticus* polyketide synthase (*pks*). There are generally two types of *pks*. The type I are large multifunctional proteins, and type II consist of several monofunctional proteins (Yu 1995).

Study of the sterigmatocystin and aflatoxin biosynthetic pathways present in *Aspergillus* sp. led to the proposition that an early step in the synthesis of these secondary metabolites must be the generation of a polyketide backbone by the polymerisation of acetate and nine malonate units by a polyketide synthase in a manner similar to fatty acid biosynthesis (Bhatnagar 1992). Research carried out on *A. parasiticus* in which the *pks* gene was disrupted demonstrated the loss of the ability to produce aflatoxins B1, B2 and G1 and of the ability to accumulate intermediates of the aflatoxin biosynthetic pathway (Feng 1995). This demonstrated conclusively that the *pks* gene product is paramount to the function of the aflatoxin biosynthetic pathway.

The *D. pini* homologue of the *A. parasiticus pks* is therefore likely to be important to the synthesis of dothistromin. Even if the dothistromin and aflatoxin/sterigmatocystin pathways diverge more than expected, the earliest steps are most likely to be conserved as polyketides form the foundation of these structures. However, melanin is also a polyketide derived secondary metabolite and the *Colletotrichum lagenarium pks*1 showed high homology to the *A. nidulans wA pks* (Takano 1995). This adds another aspect to the analysis of the dothistromin *pks* as earlier a potential melanin gene, *phn*1, was isolated from *D. pini*. It is therefore possible that the *D. pini pks* gene is required for melanin biosynthesis, dothistromin biosynthesis, or both. It is feasible that in *D. pini* the dothistromin and melanin biosynthetic pathways are not discrete, but rather share common enzymes as potentially dictated by the pathways evolution.

4.5.2.2 Polyketide Synthase Domains.

In general entire polyketide synthase genes are large (10 kb in *A. nidulans*) and contain several domains. The closest characterised homologues to the sequence generated from λBMKSA are the *Aspergillus parasiticus* and the *Aspergillus nidulans pks* genes. The *pks* encoded by each of these genes is a multifunctional novel type I polyketide synthase which has several active sites, a β-ketoacyl carrier protein synthase, an acyltransferase, duplicated acyl carrier proteins and a thioesterase. This is different to all other known *pks* genes as active sites are multiply applied to a single substrate so these genes are designated as novel type I *pks* genes (Yu 1995).

The *pks* sequence generated from the clone λBMKSA obtained from the *D. pini* genome revealed the presence of two potential *pks* domains, an acyl transferease domain and a β-ketoacyl ACP synthase domain. A pile-up (figure 4.6) of these domains demonstrated a high degree of homolgy to *A. parasiticus* with only a single residue differing in the a β-ketoacyl ACP synthase domain. The high degree of conservation present in these domains across several species indicates a probable active site.

The 2.4 kb region shown to contain these potential *pks* gene domains was present at the right hand border of the lambda clone so the remainder of the anticipated 10 kb sequence has not yet been subcloned from the genome. In a previous study of the *D. pini* lambda clone λCGV1 a potential thioesterase domain was discerned at one end (Monahan 1998). As we would expect a thioesterase domain to be present in the *pks* gene it is possible that the remaining domains lie in the genomic sequence between these two lambda clones, however the orientation of the thioesterase is wrong in λCGV1 so this is highly unlikely. The orientation of the clones relative to one another is further examined in chapter 5 (figures 5.6, 5.7 and 5.8).

The above observations coupled with the *stc*W like gene located near by lead to the tentative conclusion that perhaps this *pks* is specific to the dothistromin biosynthetic pathway and its disruption may prevent the synthesis of dothistromin and the toxic intermediates of this pathway. As well as increasing our understanding of the pathway this would provide insight as to the specific function of dothistromin (toxin or elicitor). As the *pks* gene is expected to function early in the dothistromin biosynthetic pathway it

would be a good choice for a gene disruption and could potentially lead to a novel means by which dothistroma blight could be controlled.

5. THE 'GENE CLUSTER' SO FAR.

5.1 Introduction.

The project to identify putative dothistromin biosynthetic genes was initiated under the assumption that the pathway would process intermediates in a manner similar to the aflatoxin biosynthetic pathway as the end products contain similar functional groups. Further investigation also demonstrated the presence of similar pathway intermediates.

Assuming the biosynthetic pathways were similar, a key biosynthetic gene of the aflatoxin pathway was selected, *A. parasiticus ver*1. *Ver*1 encodes a ketoreductase involved in the conversion of versicolorin A to sterigmatocystin. A probe for this gene was obtained and used to screen a *D. pini* genomic library. This screening resulted in the identification of two lambda clones, λCGV1 and λCGV2. These two lambda clones both hybridised to the *A. parasiticus ver*1 probe but fragments did not cross hybridise to each other (Gillman 1996). λCGV1 was sequenced and found to contain the *ver*-1 like *dkr*1 and other putative dothistromin biosynthetic genes (Monahan 1998). λCGV2 was also found to contain a *ver*-1 like gene, however this was determined to be more closely related to the melanin biosynthetic equivalent (*phn*1) than to the aflatoxin *ver*-1 (See chapter 3).

Due to the complexity of the biosynthetic pathway and the polyketide backbone present in both aflatoxin and dothistromin an attempt was made to isolate a homologue to the

aflatoxin *pks*. The genomic screening was carried out using an *A. parasiticus* polyketide synthase probe. This screening resulted in the isolation of the clone λBMKSA (Morgan 1997). This lambda clone contained an open reading frame which contained two putative *pks* domains, as well as a region with a 72% similarity to the *A. nidulans stc*W.

The aflatoxin/sterigmatocystin producing species of *Aspergillus* all demonstrate gene clustering. In *A. nidulans* a 60 kb region contains 25 genes which are induced only under sterigmatocystin producing conditions (Brown 1996). This region is believed to characterise most, if not all, of the genes required for the production of sterigmatocystin.

As the aflatoxin and dothistromin biosynthetic pathways demonstrate a high degree of similarity in many other features, and because both *Aspergillus* sp. and *D. pini* are ascomycetous filamentous fungi, it was anticipated that a degree of gene clustering is present in *D. pini*. This is supported by gene clustering in both the characterised clones λCGV1 (Monahan 1998) and λBMKSA. If the genes on these clones are involved in the same pathway the clones may be clustered together in the genome. In this case either the lambda clones isolated so far are overlapping (previous analysis discounted this), or are closely linked (in which case hybridisation to a Southern blot containing restriction digested *D. pini* genomic DNA would reveal a large common fragment to which all of the clones hybridised). The clones were previously compared at the level of restriction digestion (4.2.1) and no common restriction fragments were apparent. This suggests that the clones do not significantly overlap, however they may be spread throughout a cluster.

Therefore a Southern blot of *D. pini* genomic DNA was required to determine the presence of common regions between the lambda clones.

5.2 Southern Comparison Of The Lambda Clones.

5.2.1 The Southern Blot.

The digests performed for this Southern blot all contained 2 μg of genomic *D. pini* DNA, and due to previously cited *D. pini* digestion anomalies (Morgan 1997), the DNA was digested overnight. The restriction enzymes selected for this comparison were *Bam*HI, *Eco*RI, *Sal*I, *Hin*dIII, *Xba*I and *Xho*I. These enzymes were selected because previous digests of *D. pini* lambda clones and genomic DNA showed them to yield useful molecular weight fragments. Furthermore, *Xba*I was included as a previous study revealed a large (approximately 18 kb) *Xba*I fragment which could potentially unite the clones into a cluster (Monahan 1998). The digests were then Southern blotted (Section 2.13.1).

5.2.2 Hybridisation.

To determine whether the clones are derived from closely linked regions of the genome, hybridisation to the ends of the characterised lambda clones was required as sequence analysis had not revealed any linkage. The blots were probed using [α-^{32}P] labelled DNA fragments. The probes selected for this comparison were the subcloned restriction

fragments nearest to both the left and right borders of each of the three characterised lambda clones. If two clones are closely linked their end fragments should hybridise to the same genomic fragment (if the genome is cut with an enzyme which does not cut between the two clones). The restriction fragments used as probes were the λBMKSA 2.4 kb *Bam*HI/*Eco*RI and 3.0 kb *Bam*HI, the λCGV2 1.0 kb *Bam*HI and 4.4 kb *Bam* HI, and the λCGV1 2.1 kb *Eco*RI. The opposing λCGV1 2.7 kb *Eco*RI probe hybridisation was previously published (Monahan, 1998).

5.3 Chapter Five Results and Discussion.

The results of these hybridisations are depicted in table 5.1 where all of the bands which contribute to the argument that the lambda clones are clustered are highlighted in blue. The autoradiographs of these hybridisations are shown in figures 5.1-5.5. There do not appear to be any incomplete digestions and background was acceptable in all cases. However the width of the bands leads to potential fragment sizing errors.

The bands hybridised by the ends of the lambda clones which were used as probes were sized and compared. A potential junction was determined wherever a common genomic fragment was hybridised by two lambda clone ends. When the ends of the lambda clones were aligned (Figures 5.6 and 5.7) commonly hybridised genomic fragments at the junctions were identified and matched to the restriction maps. Large fragments were fitted to the diagram first as these are more likely to overlap two clones than the smaller

Figure 5.1 Southern Blotted *D. pini* Genomic DNA Restriction Profile Probed With λCGV1 2.1 kb *Eco*RI Fragment.

Lane.	**Restriction Enzyme.**	**Hybridised Band.**
1	Lambda *Eco*RI/*Hind*III	Ladder
2	Empty	Empty
3	*Bam*HI	7.8 kb
4	*Eco*RI	2.1 kb
5	*Bam*HI/*Eco*RI	2.1 kb
6	*Sal*I	11.2 kb
7	*Hin*dIII	4.9 kb /3.0 kb
8	*Xba*I	12.9 kb
9	*Xho*I	8.1 kb
10	Empty	Empty
11	Lambda *Eco*RI/*Hind*III	Ladder
12	1 kb Ladder	1 kb Ladder

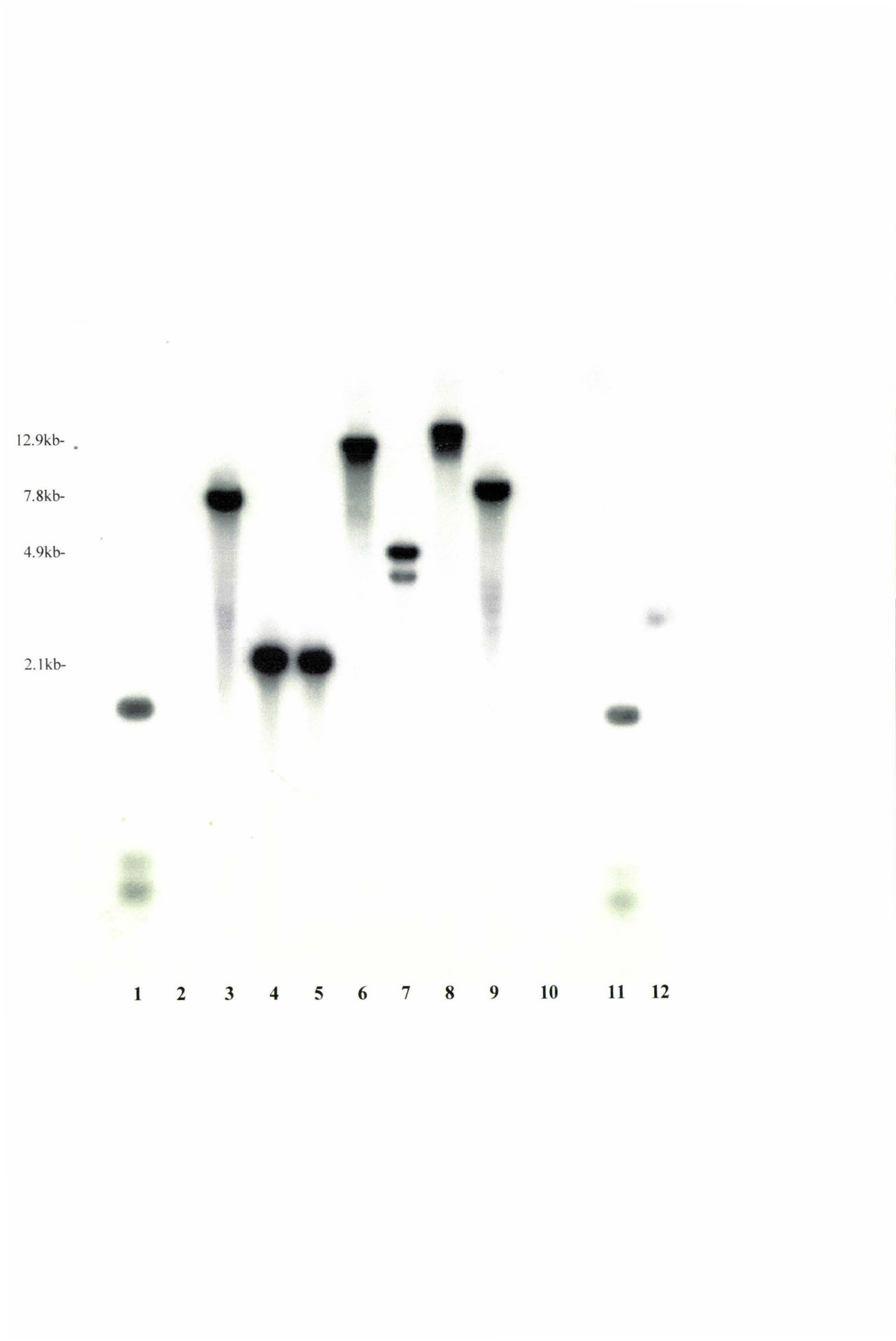

12.9kb-
7.8kb-
4.9kb-
2.1kb-
1 2 3 4 5 6 7 8 9 10 11 12

Figure 5.2 Southern Blotted *D. pini* Genomic DNA Restriction Profile Probed With λCGV2 4.4 kb BamHI Fragment.

Lane.	Restriction Enzyme.	Hybridised Band.
1	*Bam*HI	10.6 kb /2.9 kb
2	*Eco*RI	15.7 kb
3	*Bam*HI/*Eco*RI	10.6 kb
4	*Sal*I	2.9 kb /2.1 kb
5	*Hin*dIII	7.5 kb
6	*Xba*I	17.1 kb
7	*Xho*I	4.2 kb /0.6 kb

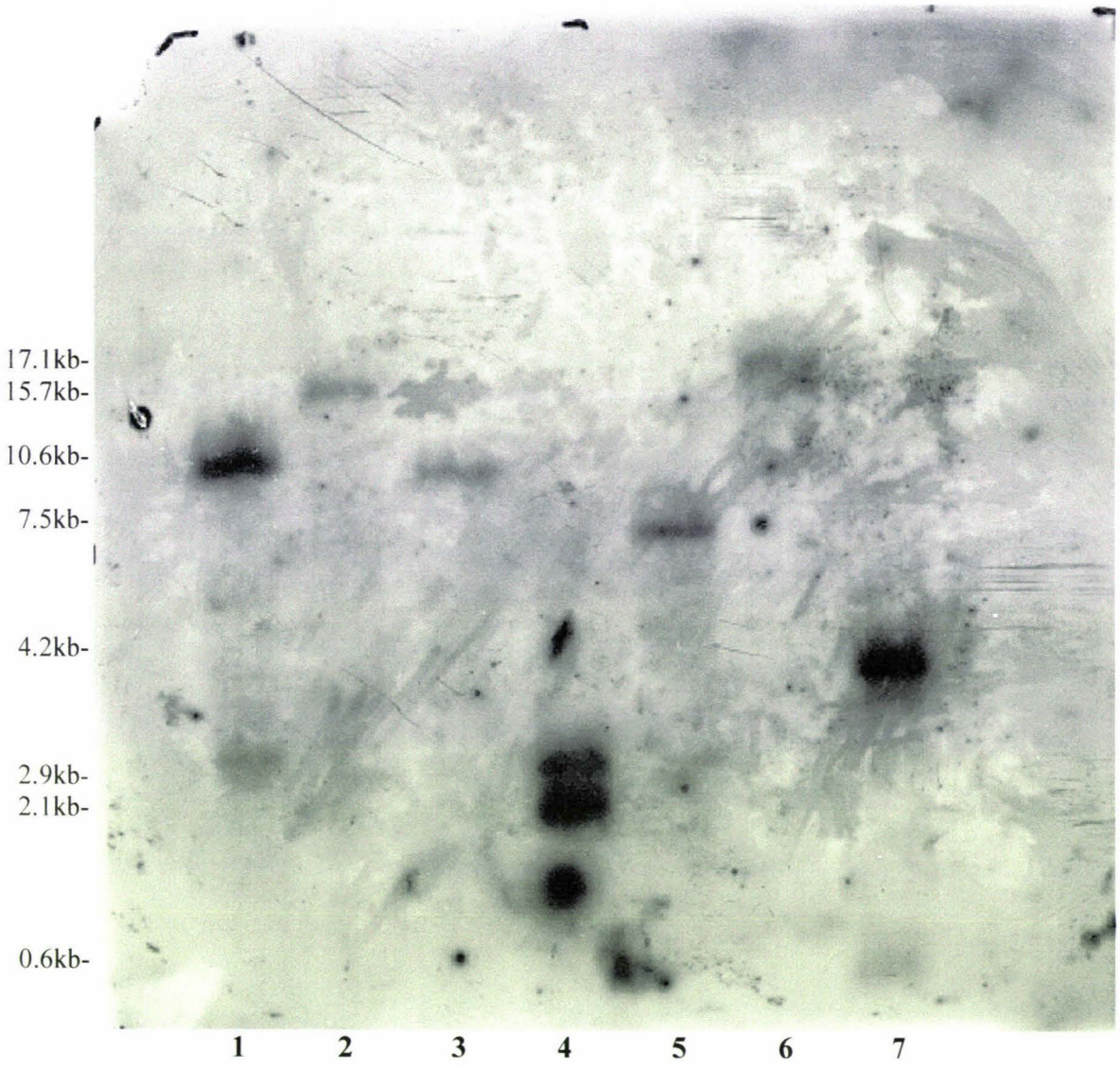

17.1kb-
15.7kb-
10.6kb-
7.5kb-
4.2kb-
2.9kb-
2.1kb-
0.6kb-
1
2
3
4
5
6
7

Figure 5.3 Southern Blotted *D. pini* Genomic DNA Restriction Profile Probed With λCGV2 1.0 kb *Bam*HI Fragment.

Lane.	Restriction Enzyme.	Hybridised Band.
1	Lambda *Eco*RI/*Hind*III	Ladder
2	*Bam*HI	1.0 kb
3	*Eco*RI	5.9 kb
4	*Bam*HI/*Eco*RI	1.0 kb
5	*Sal*I	4.6 kb
6	*Hin*dIII	5.2 kb
7	*Xba*I	9.2 kb
8	*Xho*I	4.6 kb
9	Empty	Empty
10	Lambda *Eco*RI/*Hind*III	Ladder

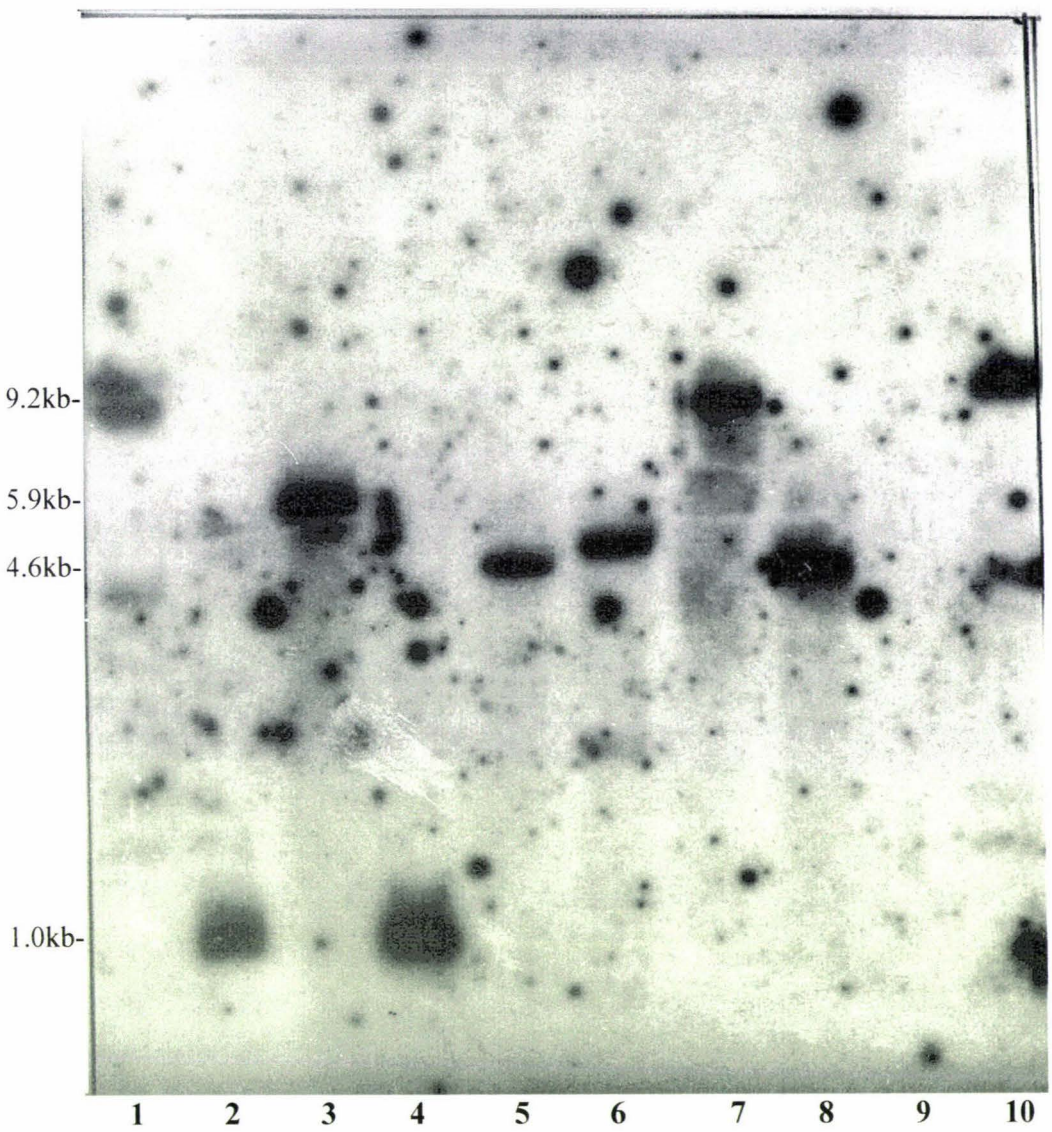

9.2kb-
5.9kb-
4.6kb-
1.0kb-
1
2
3
4
5
6
7
8
9
10

Figure 5.4 Southern Blotted *D. pini* Genomic DNA Restriction Profile Probed With λBMKSA 2.4 kb BamHI/*Eco*RI Fragment.

Lane.	Restriction Enzyme.	Hybridised Band.
1	Lambda *Eco*RI/*Hin*dIII	Ladder
2	*Bam*HI	6.3 kb
3	*Eco*RI	5.9 kb
4	*Bam*HI/*Eco*RI	2.4 kb
5	*Sal*I	4.1 kb /1.1 kb
6	*Hin*dIII	3.3 kb
7	*Xba*I	9.3 kb /2.2 kb
8	*Xho*I	2.9 kb /5.9 kb
9	Empty	Empty
10	Lambda *Eco*RI/*Hin*dIII	Ladder

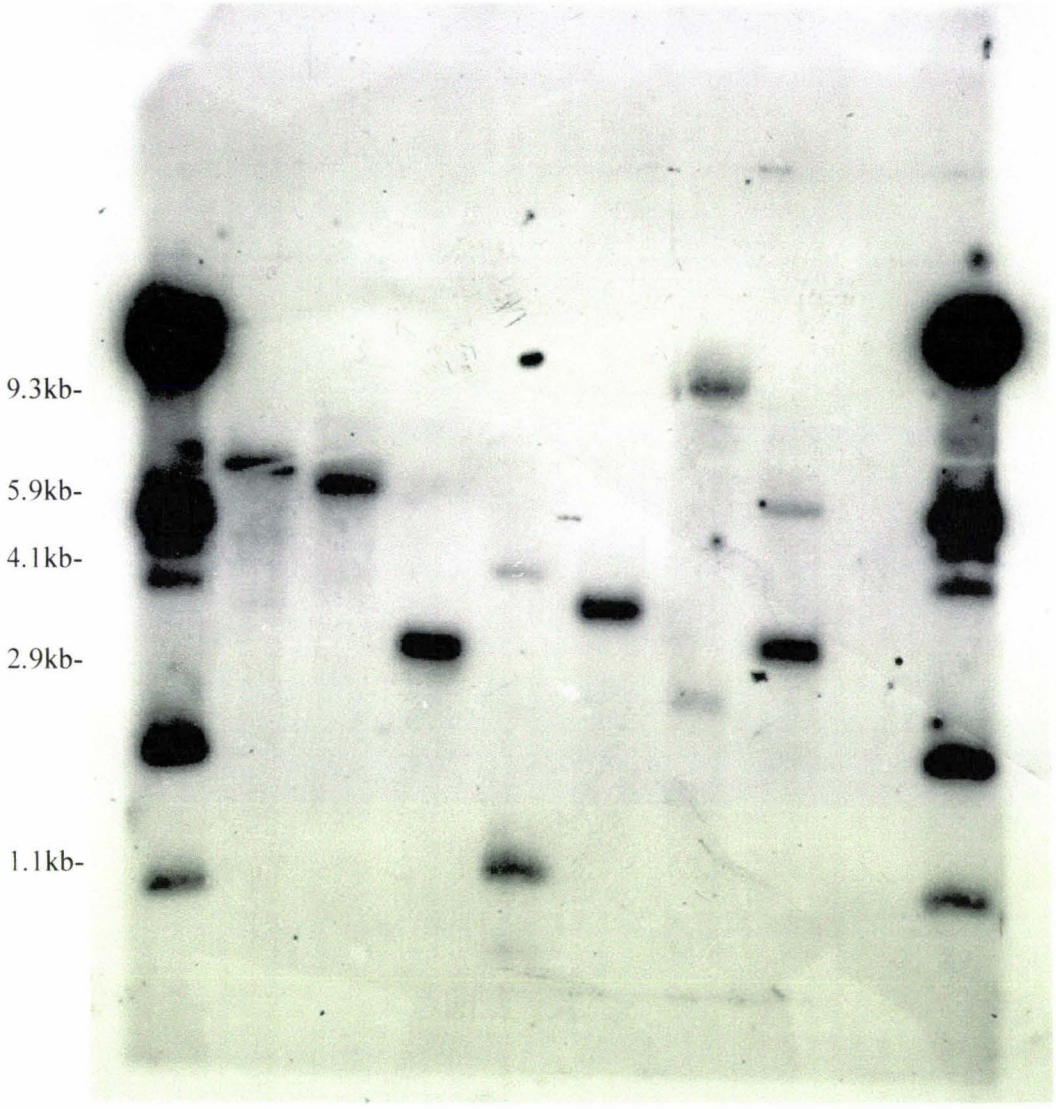

9.3kb-
5.9kb-
4.1kb-
2.9kb-
1.1kb-
1 2 3 4 5 6 7 8 9 10

Figure 5.5 Southern Blotted *D. pini* Genomic DNA Restriction Profile Probed With λBMKSA 3.0 kb BamHI Fragment.

Lane.	Restriction Enzyme.	Hybridised Band.
1	Lambda *Eco*RI/*Hin*dIII	Ladder
2	*Bam*HI	3.0 kb
3	*Eco*RI	6.8 kb
4	*Bam*HI/*Eco*RI	3.0 kb
5	*Sal*I	6.4 kb
6	*Hin*dIII	1.0 kb /1.2 kb
7	*Xba*I	7.9 kb
8	*Xho*I	5.3 kb
9	Empty	Empty
10	Lambda *Eco*RI/*Hin*dIII	Ladder

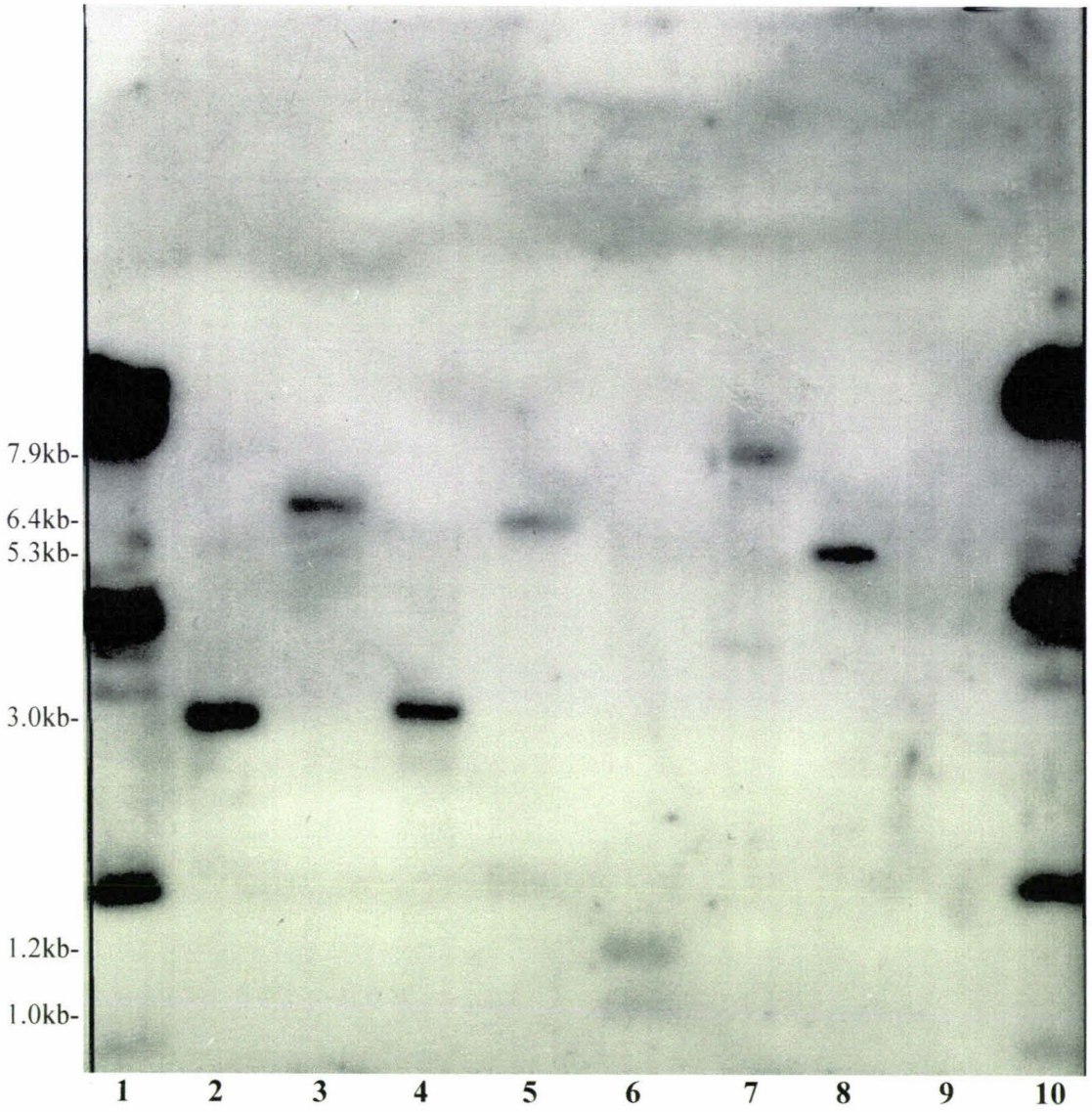

7.9kb-
6.4kb-
5.3kb-
3.0kb-
1.2kb-
1.0kb-
1
2
3
4
5
6
7
8
9
10

Table 5.1 Southern Blotting Summary.

	Fragment of λ Clone Used as Probe.					
Restriction Endonuclease. (genomic DNA digest)	λCGV1 2.1 kb *Eco*RI	*λCGV1 2.7 kb *Eco*RI	λCGV2 4.4 kb *Bam*HI	λCGV2 1.0 kb *Bam*HI	λBMKSA 2.4 kb *Bam*HI/ *Eco*RI	λBMKSA 3.0 kb *Bam*HI
*Bam*HI	7.8 kb	1.5 kb 9.9 kb	10.6 kb 2.9 kb	1.0 kb	6.3 kb	3.0 kb
*Eco*RI	2.1 kb	NT	15.7 kb	5.9 kb	5.9 kb	6.8 kb
*Bam*HI+ *Eco*RI	2.1 kb	NT	10.6 kb	1.0 kb	2.4 kb	3.0 kb
*Sal*I	11.2 kb	4.2 kb	2.9 kb 2.1 kb	4.6 kb	4.1 kb 1.1 kb	6.4 kb
*Hin*DIII	4.9 kb 3.0 kb	3.3 kb	7.5 kb	5.2 kb	3.3 kb	1.0 kb 1.2 kb
*Xba*I	12.9 kb	18.6 kb	17.1 kb	9.2 kb	9.3 kb 2.2 kb	7.9 kb
*Xho*I	8.1 kb	4.2 kb	4.2 kb 0.6 kb	4.6 kb	2.9 kb 5.9 kb	5.3 kb
Source of Results.	Figure 5.1	Monahan (1998)	Figure 5.2	Figure 5.3	Figure 5.4	Figure 5.5

* Indicates results obtained from Monahan (1998).

NT Indicates not tested.

fragments. The smaller fragments and other hybridised fragments were then fitted to the diagrams.

5.3.1 Southern Analysis.

As none of the clones were shown to overlap with each other (section 4.6.1) an investigation as to whether they were located close to one another in the genome was required. This was approached by probing the Southern blot (5.2.1) with both ends of each of the lambda clones. If the clones both overlap with a common restriction fragment the same sizes of bands on the Southern blot would be hybridised by more than one probe. The sizes of these bands could then be used to determine approximately how far apart each of the lambda clones were.

Common hybridisation bands which suggest linkage of the lambda clones λBMKSA, λCGV1 and λCGV2 are highlighted blue in table 5.1 and blue in figures 5.6 and 5.7.

The Southern blots did not reveal any large fragments which hybridised to all of the clones. However, two potential junctions became apparent, one linking λBMKSA with λCGV2 and another linking λCGV2 and λCGV1.

An alignment of the λBMKSA clone border which contained the 2.4 kb *Bam*HI/*Eco*RI probe and the λCGV2 border containing the 1.0 kb *Bam*HI probe revealed common

hybridisation signals. In the lane containing *Eco*RI digested *D. pini* genomic DNA there was a hybridisation signal which corresponds to an *Eco*RI fragment of approximately 5.9 kb in both the 2.4 kb *Bam*HI/*Eco*RI probed filter (see figure 5.4 lane 3) and the 1.0 kb *Bam*HI probed filter (see figure 5.3 lane 3). Figure 5.6 illustrates a potential origin for this fragment of genomic DNA and how it unites the two lambda clones.

A second potential point to demonstrate that these two clones are contiguous was present in the hybridisations to *Xba*I digested *D. pini* genomic DNA. The 2.4 kb *Bam*HI/*Eco*RI probe hybridised to a 9.3 kb band (figure 5.4, lane 7) and the 1.0 kb *Bam*HI probe hybridised to an equivalent region (9.2 kb) (figure 5.3, lane 7). No *Xba*I restriction sites have been identified so far in the restriction maps of either of the lambda clones being compared However, the absence of any restriction site within the fully sequenced clone λCGV1 allows an approximate position to be determined and this position is shown in figure 5.6. Regardless of position the fragment easily spans the intervening region between the lambda clones.

Further, a comparison of the lanes containing *Sal*I digested *D. pini* genomic DNA revealed bands of 4.1 kb (2.4 kb *Bam*HI/*Eco*RI probe, figure 5.4 lane 5) and 4.6 kb (1.0 kb *Bam*HI probe, figure 5.3 lane 5). These fragments are similar enough in size to be considered identical for the purposes of Southern blot comparison. The location of this band is illustrated in figure 5.6 where it is highlighted in blue as it further supports the contiguity of the lambda clones.

Figure 5.6 The Junction of λBMKSA and λCGV2.

The known restriction sites are labelled as follows:

B =*Bam*HI

E =*Eco*RI

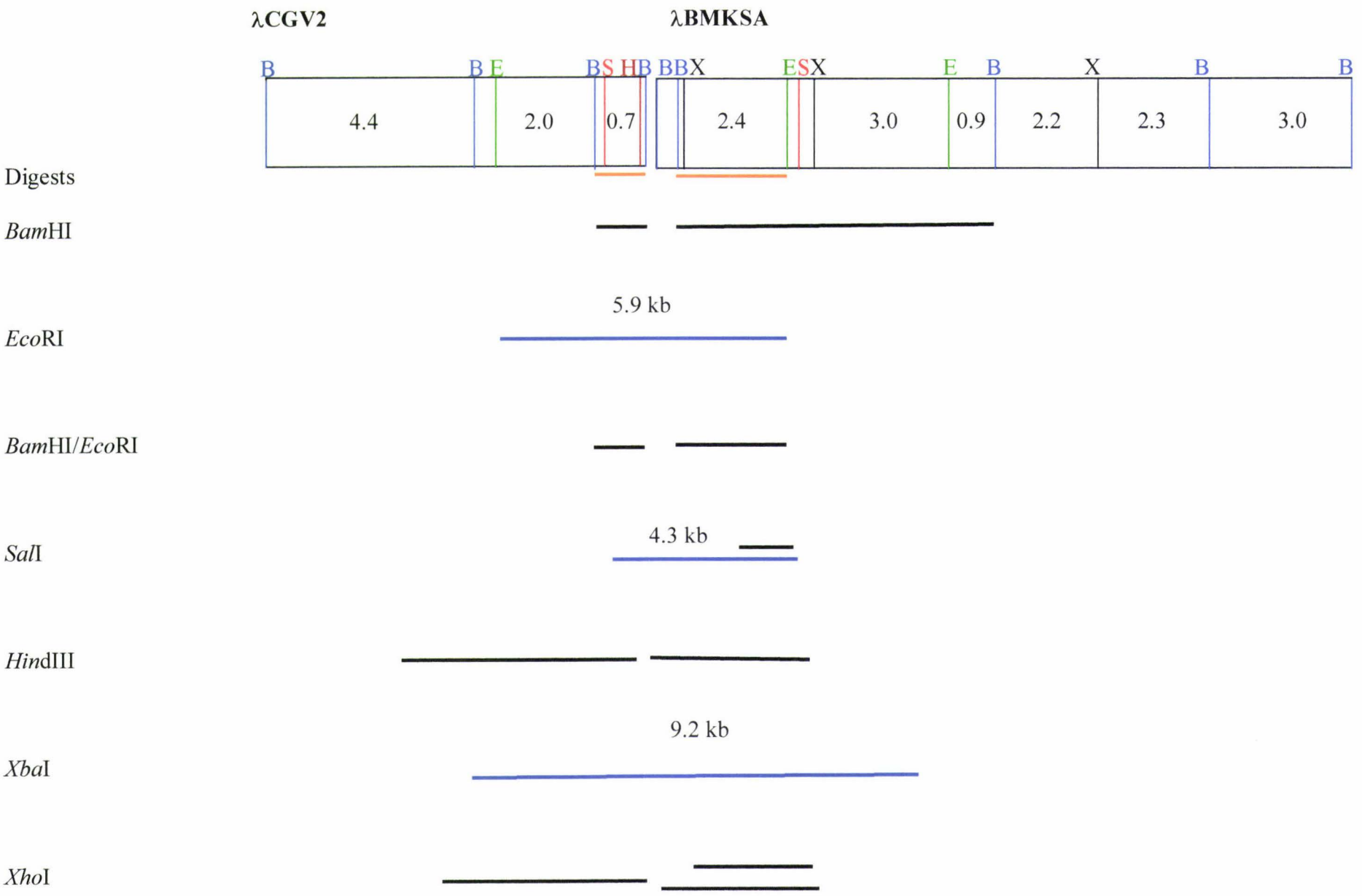
S =*Sal*I

H =*Hin*dIII

X =*Xho*I

The probes used to hybridise the fragments shown below them are labelled in orange (0.7 kb *Bam*HI and 2.4 kb *Eco*RI).

Hybridised fragments supporting contiguity of the lambda clones are represented in blue.


λCGV2
λBMKSA
B
B E
B S H B
B B X
E S X
E
B
X
B
B
4.4
2.0
0.7
2.4
3.0
0.9
2.2
2.3
3.0
Digests
BamHI
5.9 kb
EcoRI
BamHI/EcoRI
4.3 kb
SalI
HindIII
9.2 kb
XbaI
XhoI

Not all of the restriction sites have been mapped in λCGV2 (for example *Xho*I, *Xba*I, or any sites wthin the 4.4 kb *Bam*HI fragment). In turn some of the fragment positions on the lambda clone alignment are speculative, such as the position of the large *Xba*I fragment which bridges the λCGV1 and λCGV2 clones, or the position of the *Xho*I fragment relative to λCGV2 in which no restriction sites have been mapped.

The two *Xho*I fragments (2.9 kb and 5.9 kb, figure 5.6) may be attributed to incomplete restriction digestion. However the small *Sal*I fragment (4.1 kb and 1.1 kb, figure 5.6) hybridised does not appear to fit into the restriction map in any way so may be attributed to cross hybridisation to a repetitive DNA element.

The sum of the above observations lead us to the hypothesis that the λBMKSA and λCGV2 clones are contiguous. Furthermore the results consistently predict that the ends of these two lambda clones lie within 1 kb of one another. This intervening region could be generated by PCR and the primers which could be used are ver2ep1R and R156SL-pks3.

The other potential junction generated by these results was between the 4.4 kb *Bam*HI end of the λCGV2 clone and the *Eco*RI 2.7 kb end of the λCGV1 clone. This junction was first suspected due to the presence of the common *Xba*I fragment of approximately 18 kb which hybridised to both of the probes (figure 5.2, lane 6 and Monahan 1998). As

Figure 5.7 The Junction of λCGV1 and λCGV2.

The known restriction sites are labelled as follows:

B =*Bam*HI

E =*Eco*RI

S =*Sal*I

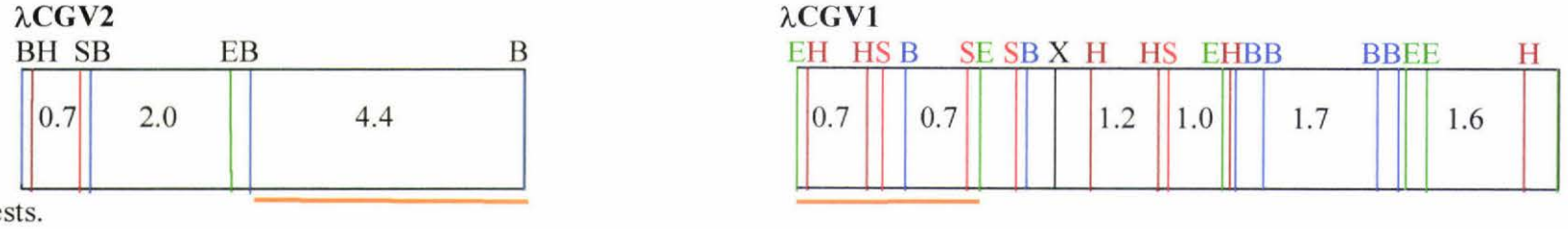
H =*Hin*dIII

X =*Xho*I

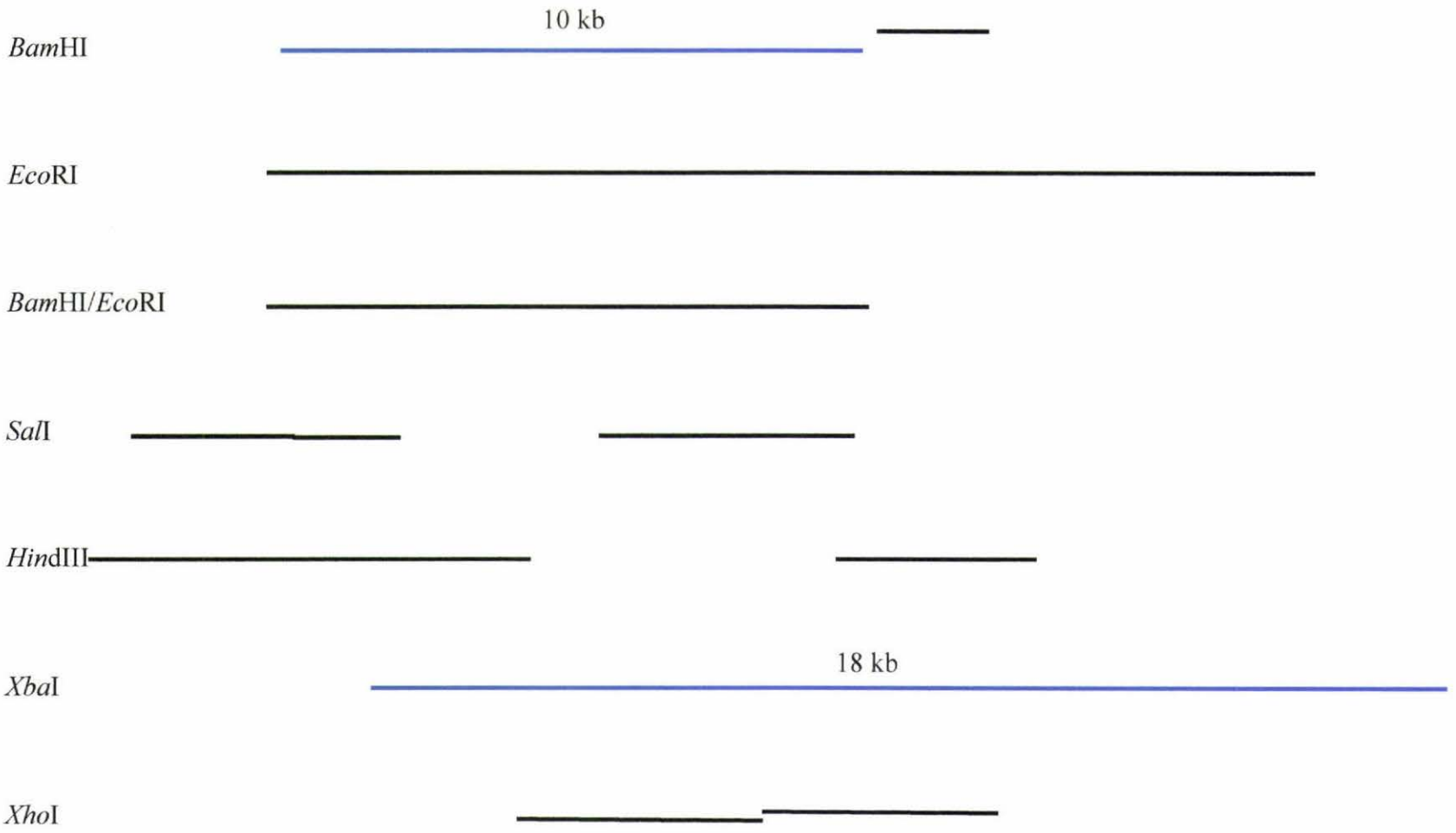
The probes used to hybridise the fragments shown below them are labelled in orange (4.4 kb *Bam*HI and 2.7 kb *Eco*RI).

Hybridised fragments supporting contiguity of the lambda clones are represented in blue.

Arrows represent the direction of large fragments beyond the scope of the diagram.


λCGV2
BH SB EB B
0.7
2.0
4.4
λCGV1
EH HS B SE SB X H HS EHBB BBEE H
0.7
0.7
1.2
1.0
1.7
1.6

Digests.


10 kb
*Bam*HI
*Eco*RI
*Bam*HI/*Eco*RI
*Sal*I
*Hin*dIII
18 kb
*Xba*I
*Xho*I

there were no *Xba*I sites present within the fully sequenced clone λCGV1 (Monahan 1998) it is likely that there is one present within the unsequenced and unmapped region of the 4.4 kb *Bam*HI fragment of the clone λCGV2. For the potential location of this fragment relative to the clones λCGV1 and λCGV2 see figure 5.7.

The second observation supporting the contiguity of these two clones is present in hybridisations of both of the above probes to the *Bam*HI digests. In this lane the two probes hybridised to a fragment of approximately 10 kb (see figure 5.2, lane 1 and Monahan, page 85, 1998). Figure 5.7 indicates a likely origin for the fragment which further supports the hypothesis that the two clones are near one another within the genome of *D. pini*. It is important to note that the apparent *Bam*HI restriction site situated at the right hand border of the λCGV2 clone (as drawn in figure 5.7) is contained in the vector arm, not the subcloned DNA.

The above arguments suggest that the end of the λCGV2 clone containing the 4.4 kb *Bam*HI probe and the end of the λCGV1 clone containing the 2.7 kb *Eco*RI probe are contiguous and lie within 4.2 kb of one another. This region could also be generated using PCR but first the 4.4 kb *Bam*HI λCGV2 fragment should be sequenced so that a PCR primer could be designed within this region and the PCR product will not be as large.

The *Eco*RI and EcoRI/BamHI lanes were only probed with the lambda clone λCGV2 as the other results were obtained from Monahan (1998) and these digests were not

included. The *Eco*RI and EcoRI/BamHI fragments hybridised by the 4.4 kb *Bam*HI probe do not appear to fit the restriction map, however if the genomic digests were partial these fragments may join the clones as shown in figure 5.7. This is very speculative so not included in the case for the contiguity of the clones. I believe this inconsistency sheds doubt on the λCGV1 λCGV2 junction.

A band anomaly is also present: a 2.9 kb band which hybridised to the λCGV2 4.4kb *Bam*HI probe in the *Bam*HI digest lane. This band does not appear to be an internal band of the λCGV2 clone, nor is it likely to be from the region dividing the λCGV2 clone from the λCGV1 clone as there do not appear to be any *Bam*HI restriction sites between the two (assuming the genomic *Bam*HI fragment spanning the clones exists). As this band, or one smaller, is not present in the *Bam*HI/*Eco*RI digest it suggests that this band was not relevant to the intersecting region and its presence may be attributed to it containing a sequence which is repeated elsewhere in the genome.

An other point of inaccuracy in the comparison is that no *Xba*I sites have been mapped in either the lambda clone λCGV2 (data suggests none present) or the lambda clone λCGV1. This leads to the positions of these hybridised fragments being speculative.

However, Southern blotting comparisons are inherently inaccurate due to the methods used to disperse and size the DNA, and restriction digestion can generate fragments of similar sizes from all areas of the genome. In addition as two copies of *ver*-1 (cf *dkr*1 and *phn*1) have been identified we would expect some cross hybridisation between other

regions of these clusters, so this result can not be accepted alone. Further experimentation will be required to confirm these results.

The overall picture developed by the sum of these observations is that the end of the clone λBMKSA (containing a 2.4 kb *Bam*HI/*Eco*RI fragment) lies next to the end of the clone λCGV2 (with a 1.0 kb *Bam*HI fragment) and that the distance between the clones is less than 1 kb. In turn the other end of the clone λCGV2 borders the end of the clone λCGV1 (which contains an *Eco*RI 2.7 kb fragment) with an intervening region of less than 4.2 kb, however the evidence in this case is not as strong. This suggests that the three clones are in fact clustered and could all potentially contain dothistromin biosynthetic genes (figure 5.8).

The relationship between the λCGV1, λCGV2 and λBMKSA clones may be elucidated by a simple PCR reaction, the product of which could be subcloned and sequenced to provide the intervening sequence without requiring further library screening. The next obvious step would be to fully sequence the lambda clones λCGV1, λCGV2 and λBMKSA as initial sequence of λCGV1 and λBMKSA revealed potential dothistromin biosynthetic genes and more of the cluster may be contained in λCGV2. If the gene density of λCGV1 is representative of the region, sequencing of λBMKSA and λCGV2 is likely to uncover more genes within this region.

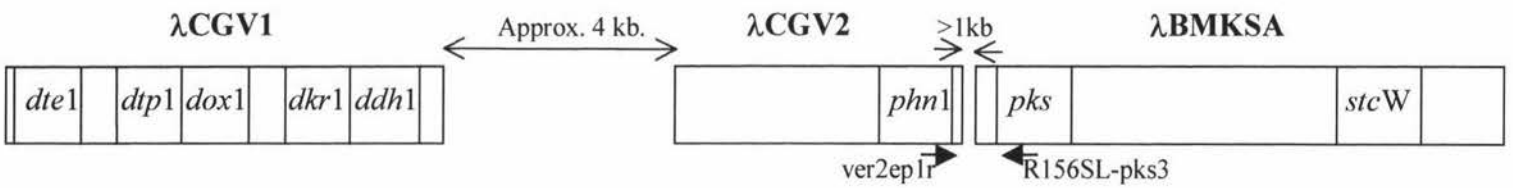
After the lambda clones already being studied have been fully characterised the following step would be to extend the ends of the contig so far developed by chromosome walking.

Figure 5.8 The Distribution of Three lambda Clones in a Putative Dothistromin Biosynthetic Cluster.

The possible arrangement of the lambda clones in the genome. Clone dimensions are scaled relative to one another, however genes are labelled relative to location not size.

Putative dothistromin Biosynthetic genes are shown in abbreviated form on the lambda clones.

The arrows represent the primers required in a PCR reaction to confirm the order of the lambda clones λCGV2 and λBMKSA and their association to one another.


λCGV1
Approx. 4 kb.
λCGV2
>1kb
λBMKSA
dte1
dtp1
dox1
dkr1
ddh1
phn1
pks
stcW
ver2ep1r
R156SL-pks3

This would hopefully reveal more of the genes in this putative dothistromin biosynthetic cluster.

However, now that there is a characterised region containing clustered genes, it is essential to confirm the function of at least some of these genes prior to the expenditure required for large scale sequencing. Specific gene disruptions by transformational knockout of genes believed to be paramount to dothistromin biosynthesis such as *dkr*1 (and *phn*1 as its similarity indicates it may catalyse the same reaction) or more likely *pks* should be carried out. The results of these disruptions should then govern the next stage of the project.

CONCLUSIONS AND FUTURE DIRECTIONS.

The region of the lambda clone λCGV2 believed to contain a second copy of *dkr*1 (aflatoxin *ver*-1 equivalent) was sequenced and analysis of this region suggested that the gene contained was more similar to a melanin biosynthetic *phn*1 gene. This led to the premature conclusion that the remainder of the clone was unlikely to contain further relevant genes. Similarly, the *pks* containing λBMKSA was analysed to demonstrate that as well as a potential dothistromin biosynthetic *pks* it also contained a gene with a high degree of similarity to the aflatoxin *stc*W gene. The presence of these t wo genes, both potentially relevant to dothistromin biosynthesis suggests that this clone may be clustered with the clone λCGV1 which was also shown to contain clustered genes potentially relevant to dothistromin biosynthesis.

To determine whether the genes were clustered a Southern blot of restriction digested *D. pini* DNA was probed with the end fragments of each lambda clone. If the clones were located within a restriction site of one another the probes would hybridise to a common genomic restriction fragment and give us an estimate as to how far apart the lambda clones were. The results of this analysis suggested that the borders of the lambda clones neighboured one another.

As the three lambda clones were shown to be clustered in the genome, with the *phn*1 containing λCGV2 in between the two shown to contain putative dothistromin biosynthetic genes, it follows that there is potentially a cluster of dothistromin biosynthetic genes.

All of the research carried out on the dothistromin biosynthetic pathway has been based on sequence similarity to the model aflatoxin biosynthetic pathway. The project has identified a potential cluster of putative dothistromin biosynthetic genes, and before any further work is carried out gene disruptions must be performed to confirm the relevance of the genes so far isolated. This should lead to several dothistromin minus mutants (one of the aims of the dothistromin project) and allow an accurate determination of the relevance of dothistromin to the pathogenicity of *D. pini*.

Once the gene disruptions have been carried out, assuming dothistromin minus mutants are produced, further sequence analysis should be carried out. This should include sequencing all of the subcloned lambda clone fragments to isolate further genes. Assuming that the gene density across the potential cluster is similar to that across the fully sequenced λCGV1 this should uncover several further genes.

While this sequencing is being carried out PCR should be used to confirm the locations of the lambda clones in the genome and the PCR products should be cloned to provide the sequence between the lambda clones.

I suspect that not all of the genes located are relevant to dothistromin biosynthesis. The location of the *phn*1 gene, unlikely to be involved in dothistromin biosynthesis, in the midst of the cluster suggests that perhaps the dothistromin biosynthetic cluster and a melanin biosynthetic cluster are located near one another and perhaps some genes are involved in both pathways.

Appendix 1: DNA Sequences.

1.1 λCGV2 Sequence (2.0 kb *Bam*HI/*Eco*RI fragment and 1.0 kb *Bam*HI fragment).

1.2 λBMKSA Sequences.

1.2.1 R156 Sequence.

1.2.2 R161 partial sequence generated using pUCfwd primer.

1.2.3 R162 partial sequence generated using pUCfwd primer.

1.2.4 R162 partial sequence generated using pUCrev primer.

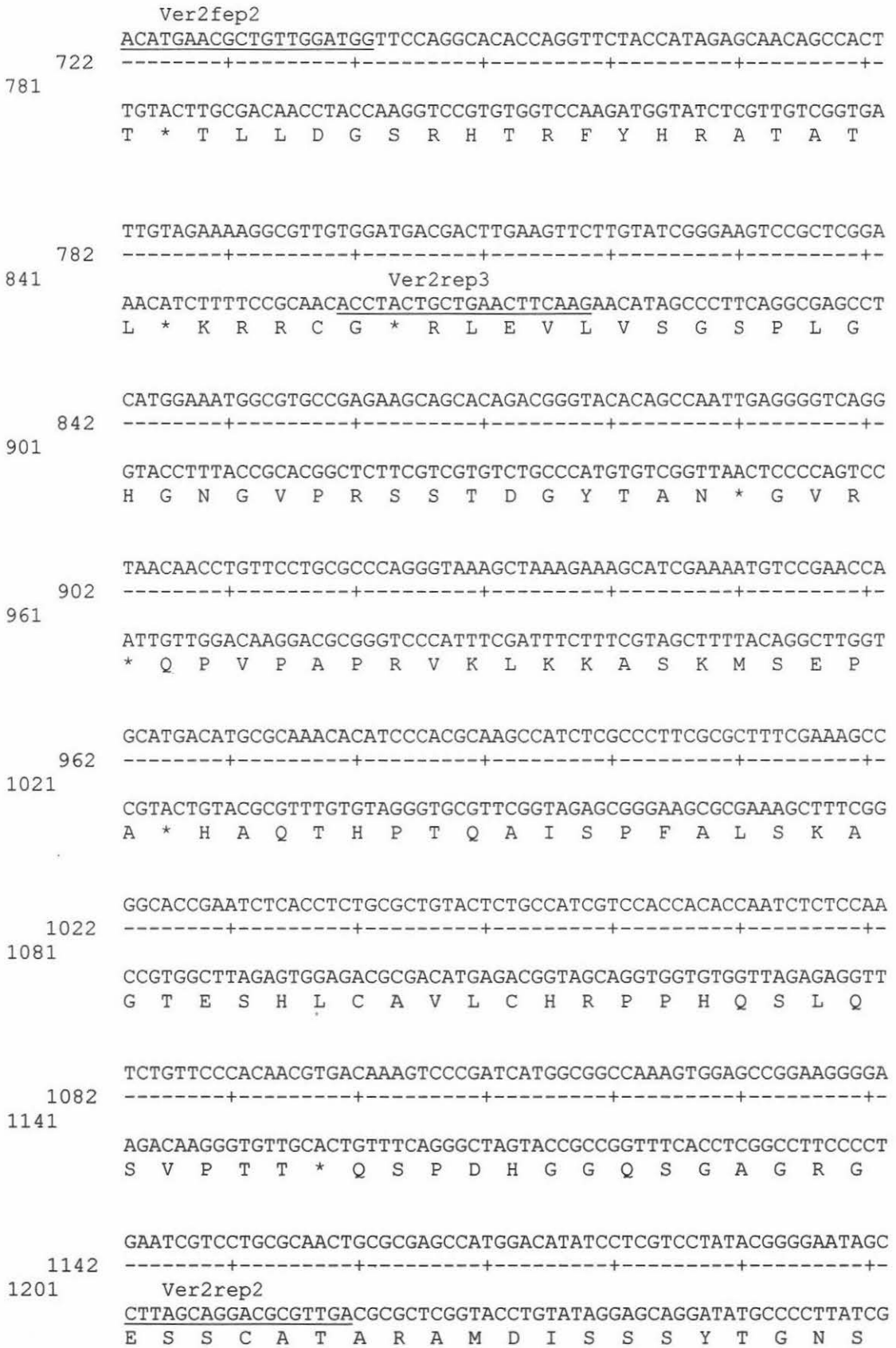
1.2.5 R163 partial sequence generated using pUCfwd primer.

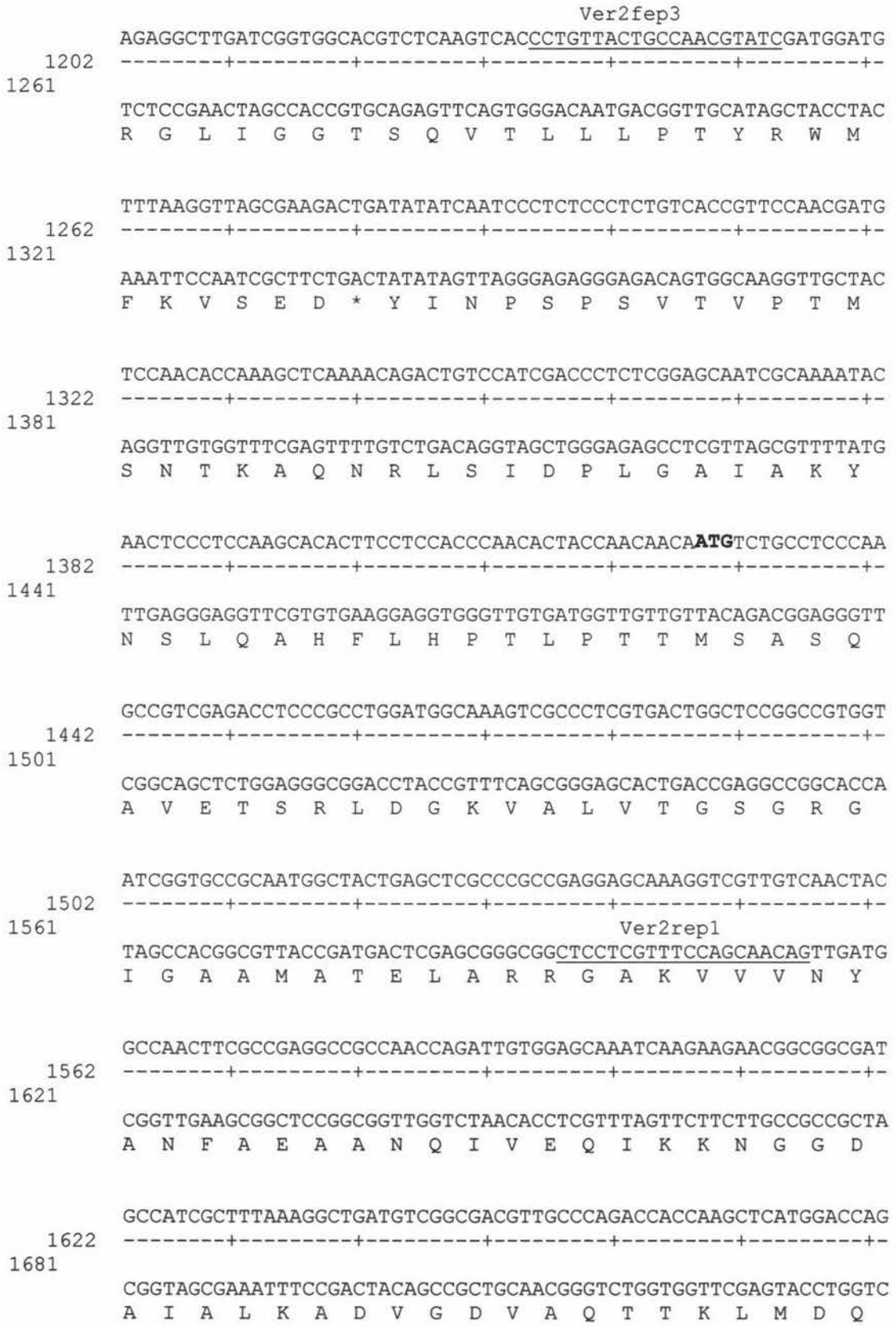
1.1 λCGV2 Sequence.

(Primers Uderlined)

```
        AATCCATGTCCGTACGAGGTTGTTTGCGGCTTGCAGTAGCATATGATCTATGCCATAGAG
      2 --------+---------+---------+---------+---------+---------+-
61
        TTAGGTACAGGCATGCTCCAACAAACGCCGAACGTCATCGTATACTAGATACGGTATCTC
        N  P  C  P  Y  E  V  V  C  G  L  Q  *  H  M  I  Y  A  I  E


        GCAGCCCAAGCGCTGATCTCTTCTGCAATATCGGCATCGCAATGGGACGAAATCGCCTGA
     62 --------+---------+---------+---------+---------+---------+-
121
        CGTCGGGTTCGCGACTAGAGAAGACGTTATAGCCGTAGCGTTACCCTGCTTTAGCGGACT
        A  A  Q  A  L  I  S  S  A  I  S  A  S  Q  W  D  E  I  A  *


        CTTCAGACTTGCTGGCAATGAGCAGGCTGGGAGTTGCCATTGCGAGGGATAGGCGCTAGC
    122 --------+---------+---------+---------+---------+---------+-
181
        GAAGTCTGAACGACCGTTACTCGTCCGACCCTCAACGGTAACGCTCCCTATCCGCGATCG
        L  Q  T  C  W  Q  *  A  G  W  E  L  P  L  R  G  I  G  A  S


        TGTGCGTGTCCACACACATTCTAATCAAGCCTCTGTCGCCATCTCCCGGAGACATTTTGC
    182 --------+---------+---------+---------+---------+---------+-
241
        ACACGCACAGGTGTGTGTAAGATTAGTTCGGAGACAGCGGTAGAGGGCCTCTGTAAAACG
        C  A  C  P  H  T  F  *  S  S  L  C  R  H  L  P  E  T  F  C
```

```
        CTTCTTGATTCTCGATGAAGGTAGCGATATCAAAGAAAAGATATTCTACAAGTCGTCCTG
    242 --------+---------+---------+---------+---------+---------+-
301
        GAAGAACTAAGAGCTACTTCCATCGCTATAGTTTCTTTTCTATAAGATGTTCAGCAGGAC
        L  L  D  S  R  *  R  *  R  Y  Q  R  K  D  I  L  Q  V  V  L

        GGTGAGTATGTTCTAATTGCGTGCTGTTGCTACAAGATACAAGTTGTCTGTTCGCTGTAC
    302 --------+---------+---------+---------+---------+---------+-
361                                                      Ver2rep4
        CCACTCATACAAGATTAACGCACGACAACGATGTTCTATGTTCAACAGACAAGCGACATG
        G  E  Y  V  L  I  A  C  C  C  Y  K  I  Q  V  V  C  S  L  Y

           Ver2fep1
        CAAACATCAGCCACAGCAAGAACGAGCTACCCTCTGAAAGTGCGGCGGTCGTGGGCGTCT
    362 --------+---------+---------+---------+---------+---------+-
421
        GTTTGTAGTCGGTGTCGTTCTTGCTCGATGGGAGACTTTCACGCCGCCAGCACCCGCAGA
        Q  T  S  A  T  A  R  T  S  Y  P  L  K  V  R  R  S  W  A  S

        TAGATCAGGCCATCAGCGTCCTGGACCTCATGCTCACTATCCTCAACCAACACTCGATCG
    422 --------+---------+---------+---------+---------+---------+-
481
        ATCTAGTCCGGTAGTCGCAGGACCTGGAGTACGAGTGATAGGAGTTGGTTGTGAGCTAGC
        *  I  R  P  S  A  S  W  T  S  C  S  L  S  S  T  N  T  R  S

        GCACCGAGCTGACATCGCCGCTGCGACTGTATGAACGAAGCCGAGGCCATACCATGATTG
    482 --------+---------+---------+---------+---------+---------+-
541
        CGTGGCTCGACTGTAGCGGCGACGCTGACATACTTGCTTCGGCTCCGGTATGGTACTAAC
        A  P  S  *  H  R  R  C  D  C  M  N  E  A  E  A  I  P  *  L

        CCACAATTCATTGACTGCAGATCACGGTAGATGGGTGGTCGCTGAAGGAAATTTTGGGTT
    542 --------+---------+---------+---------+---------+---------+-
601
        GGTGTTAAGTAACTGACGTCTAGTGCCATCTACCCACCAGCGACTTCCTTTAAAACCCAA
        P  Q  F  I  D  C  R  S  R  *  M  G  G  R  *  R  K  F  W  V

        CGAGAAGCCCGTGAAACAGTGAAACAGCCAGGGATGATATCGCTCGGTGTGAGAGCTTTG
    602 --------+---------+---------+---------+---------+---------+-
661
        GCTCTTCGGGCACTTTGTCACTTTGTCGGTCCCTACTATAGCGAGCCACACTCTCGAAAC
        R  E  A  R  E  T  V  K  Q  P  G  M  I  S  L  G  V  R  A  L

        CCTTCGGGCTTCCTCCTTCAGTCACTCCAATAGTATTCGTCTTAGGATGGGAAGACGACG
    662 --------+---------+---------+---------+---------+---------+-
721
        GGAAGCCCGAAGGAGGAAGTCAGTGAGGTTATCATAAGCAGAATCCTACCCTTCTGCTGC
        P  S  G  F  L  L  Q  S  L  Q  *  Y  S  S  *  D  G  K  T  T
```


Ver2fep2
ACATGAACGCTGTTGGATGGTTCCAGGCACACCAGGTTCTACCATAGAGCAACAGCCACT
722
781
TGTACTTGCGACAACCTACCAAGGTCCGTGTGGTCCAAGATGGTATCTCGTTGTCGGTGA
T * T L L D G S R H T R F Y H R A T A T
TTGTAGAAAAGGCGTTGTGGATGACGACTTGAAGTTCTTGTATCGGGAAGTCCGCTCGGA
782
841
Ver2rep3
AACATCTTTTCCGCAACACCTACTGCTGAACTTCAAGAACATAGCCCTTCAGGCGAGCCT
L * K R R C G * R L E V L V S G S P L G
CATGGAAATGGCGTGCCGAGAAGCAGCACAGACGGGTACACAGCCAATTGAGGGGTCAGG
842
901
GTACCTTTACCGCACGGCTCTTCGTCGTGTCTGCCCATGTGTCGGTTAACTCCCCAGTCC
H G N G V P R S S T D G Y T A N * G V R
TAACAACCTGTTCCTGCGCCCAGGGTAAAGCTAAAGAAAGCATCGAAAATGTCCGAACCA
902
961
ATTGTTGGACAAGGACGCGGGTCCCATTTCGATTTCTTTCGTAGCTTTTACAGGCTTGGT
* Q P V P A P R V K L K K A S K M S E P
GCATGACATGCGCAAACACATCCCACGCAAGCCATCTCGCCCTTCGCGCTTTCGAAAGCC
962
1021
CGTACTGTACGCGTTTGTGTAGGGTGCGTTCGGTAGAGCGGGAAGCGCGAAAGCTTTCGG
A * H A Q T H P T Q A I S P F A L S K A
GGCACCGAATCTCACCTCTGCGCTGTACTCTGCCATCGTCCACCACACCAATCTCTCCAA
1022
1081
CCGTGGCTTAGAGTGGAGACGCGACATGAGACGGTAGCAGGTGGTGTGGTTAGAGAGGTT
G T E S H L C A V L C H R P P H Q S L Q
TCTGTTCCCACAACGTGACAAAGTCCCGATCATGGCGGCCAAAGTGGAGCCGGAAGGGGA
1082
1141
AGACAAGGGTGTTGCACTGTTTCAGGGCTAGTACCGCCGGTTTCACCTCGGCCTTCCCCT
S V P T T * Q S P D H G G Q S G A G R G
GAATCGTCCTGCGCAACTGCGCGAGCCATGGACATATCCTCGTCCTATACGGGGAATAGC
1142
1201
Ver2rep2
CTTAGCAGGACGCGTTGACGCGCTCGGTACCTGTATAGGAGCAGGATATGCCCCTTATCG
E S S C A T A R A M D I S S S Y T G N S


Ver2fep3
AGAGGCTTGATCGGTGGCACGTCTCAAGTCACCCTGTTACTGCCAACGTATCGATGGATG
1202
1261
TCTCCGAACTAGCCACCGTGCAGAGTTCAGTGGGACAATGACGGTTGCATAGCTACCTAC
R G L I G G T S Q V T L L L P T Y R W M
TTTAAGGTTAGCGAAGACTGATATATCAATCCCTCTCCCTCTGTCACCGTTCCAACGATG
1262
1321
AAATTCCAATCGCTTCTGACTATATAGTTAGGGAGAGGGAGACAGTGGCAAGGTTGCTAC
F K V S E D * Y I N P S P S V T V P T M
TCCAACACCAAAGCTCAAAACAGACTGTCCATCGACCCTCTCGGAGCAATCGCAAAATAC
1322
1381
AGGTTGTGGTTTCGAGTTTTGTCTGACAGGTAGCTGGGAGAGCCTCGTTAGCGTTTTATG
S N T K A Q N R L S I D P L G A I A K Y
AACTCCCTCCAAGCACACTTCCTCCACCCAACACTACCAACAACAATGTCTGCCTCCCAA
1382
1441
TTGAGGGAGGTTCGTGTGAAGGAGGTGGGTTGTGATGGTTGTTGTTACAGACGGAGGGTT
N S L Q A H F L H P T L P T T M S A S Q
GCCGTCGAGACCTCCCGCCTGGATGGCAAAGTCGCCCTCGTGACTGGCTCCGGCCGTGGT
1442
1501
CGGCAGCTCTGGAGGGCGGACCTACCGTTTCAGCGGGAGCACTGACCGAGGCCGGCACCA
A V E T S R L D G K V A L V T G S G R G
ATCGGTGCCGCAATGGCTACTGAGCTCGCCCGCCGAGGAGCAAAGGTCGTTGTCAACTAC
1502
1561
Ver2rep1
TAGCCACGGCGTTACCGATGACTCGAGCGGGCGGCTCCTCGTTTCCAGCAACAGTTGATG
I G A A M A T E L A R R G A K V V V N Y
GCCAACTTCGCCGAGGCCGCCAACCAGATTGTGGAGCAAATCAAGAAGAACGGCGGCGAT
1562
1621
CGGTTGAAGCGGCTCCGGCGGTTGGTCTAACACCTCGTTTAGTTCTTCTTGCCGCCGCTA
A N F A E A A N Q I V E Q I K K N G G D
GCCATCGCTTTAAAGGCTGATGTCGGCGACGTTGCCCAGACCACCAAGCTCATGGACCAG
1622
1681
CGGTAGCGAAATTTCCGACTACAGCCGCTGCAACGGGTCTGGTGGTTCGAGTACCTGGTC
A I A L K A D V G D V A Q T T K L M D Q

Ver2fep4

GCCGTTGAGCACTACGGCCAGCTCGACATTGTCTGCTCCAACTCCGGCGTTGTCTCTTTC

1682 --------+---------+---------+---------+---------+---------+-
1741

CGGCAACTCGTGATGCCGGTCGAGCTGTAACAGACGAGGTTGAGGCCGCAACAGAGAAAG
A V E H Y G Q L D I V C S N S G V V S F

GGCCACTTGAAGGACGTCACTGAGGAGGAGTTCGACCGTGTGTTCCGCATCAACACCCGC

1742 --------+---------+---------+---------+---------+---------+-
1801

CCGGTGAACTTCCTGCAGTGACTCCTCCTCAAGCTGGCACACAAGGCGTAGTTGTGGGCG
G H L K D V T E E E F D R V F R I N T R

GGCCAGTTCTTCGTTGCCCGTGAGGCGTACAAGCACTTGAGCGTTGGTGGTCGCATTATT

1802 --------+---------+---------+---------+---------+---------+-
1861

CCGGTCAAGAAGCAACGGGCACTCCGCATGTTCGTGAACTCGCAACCACCAGCGTAATAA
G Q F F V A R E A Y K H L S V G G R I I

CTTATGGGTTCCATCACTGGTCAGGCCAAGGGTGTGCCAAAGCACGCCGTCTACTCTGGA

1862 --------+---------+---------+---------+---------+---------+-
1921

GAATACCCAAGGTAGTGACCAGTCCGGTTCCCACACGGTTTCGTGCGGCAGATGAGACCT
L M G S I T G Q A K G V P K H A V Y S G

TCCAAGGGCGCCATCGAGACTTTCGTCCGATGCATGGCCATCGACTGTGGTGACAAGAAG

1922 --------+---------+---------+---------+---------+---------+-
1981

AGGTTCCCGCGGTAGCTCTGAAAGCAGGCTACGTACCGGTAGCTGACACCACTGTTCTTC
S K G A I E T F V R C M A I D C G D K K

ATCACTGTCAACTGCGTTGCTCCAGGTGGTATCAAGACTGACATGTACCACGCTGTATGC

1982 --------+---------+---------+---------+---------+---------+-
2041

TAGTGACAGTTGACGCAACGAGGTCCACCATAGTTCTGACTGTACATGGTGCGACATACG
I T V N C V A P G G I K T D M Y H A V C

CGCGAATACATTCCGAACGGTGAGAACCTCAGCAACGATCAGGTCGACGAGGTAAGCATA

2042 --------+---------+---------+---------+---------+---------+-
2101

GCGCTTATGTAAGGCTTGCCACTCTTGGAGTCGTTGCTAGTCCAGCTGCTCCATTCGTAT
R E Y I P N G E N L S N D Q V D E V S I

ATCCACGCTGGTGGTCATGATATACACATTGCTAATGAGCCGTCGATAGTACGCCAAGAC

2102 --------+---------+---------+---------+---------+---------+-
2161

TAGGTGCGACCACCAGTACTATATGTGTAACGATTACTCGGCAGCTATCATGCGGTTCTG
I H A G G H D I H I A N E P S I V R Q D

```
        GTGGTCTCCAATGCAGCGTGTTGGCCAGCCAATCGACATTGCCCGTGTTGTCTGCTTCCT
   2162 --------+---------+---------+---------+---------+---------+-
2221
        CACCAGAGGTTACGTCGCACAACCGGTCGGTTAGCTGTAACGGGCACAACAGACGAAGGA
        V  V  S  N  A  A  C  W  P  A  N  R  H  C  P  C  C  L  L  P

        CGCCTCCCAAGACGGCGAATGGGTCAACGGCAAGGTCATTGGCATTGATGGTGCTGCATG
   2222 --------+---------+---------+---------+---------+---------+-
2281
        GCGGAGGGTTCTGCCGCTTACCCAGTTGCCGTTCCAGTAACCGTAACTACCACGACGTAC
        R  L  P  R  R  R  M  G  Q  R  Q  G  H  W  H  *  W  C  C  M

           Ver2rp1R
        CATGTAAACGTCCAGTCCTCGTGGCTTGGAACGACTTTTGGCTGGCAGCGTGTTCACGAG
   2282 --------+---------+---------+---------+---------+---------+-
2341
        GTACATTTGCAGGTCAGGAGCACCGAACCTTGCTGAAAACCGACCGTCGCACAAGTGCTC
        H  V  N  V  Q  S  S  W  L  G  T  T  F  G  W  Q  R  V  H  E

        CATTTCGACTGACCACAGTATGGACGAAAAGATGGAATTATGGATGTCACGGCTGAAATG
   2342 --------+---------+---------+---------+---------+---------+-
2401                                                Ver2ep1F
        GTAAAGCTGACTGGTGTCATACCTGCTTTTCTACCTTAATACCTACAGTGCCGACTTTAC
        H  F  D  *  P  Q  Y  G  R  K  D  G  I  M  D  V  T  A  E  M

        GCGCAAGGCGGATGGTGTTGAGACGTTCTTCGCAAAAGATGTAGTGTCTAGACAGAGATG
   2402 --------+---------+---------+---------+---------+---------+-
2461
        CGCGTTCCGCCTACCACAACTCTGCAAGAAGCGTTTTCTACATCACAGATCTGTCTCTAC
        A  Q  G  G  W  C  *  D  V  L  R  K  R  C  S  V  *  T  E  M

        GCCGACGGTACAGAGTATCATAGATGTCCGATTCAAGTGCTGAATAATGGTCATTGTACC
   2462 --------+---------+---------+---------+---------+---------+-
2521
        CGGCTGCCATGTCTCATAGTATCTACAGGCTAAGTTCACGACTTATTACCAGTAACATGG
        A  D  G  T  E  Y  H  R  C  P  I  Q  V  L  N  N  G  H  C  T

        CAGTCACTGGAGCCGCATTCCGCCGAGCCTCTCAATGTTCTGACATCTTTGTTGCCTGAG
   2522 --------+---------+---------+---------+---------+---------+-
2581
        GTCAGTGACCTCGGCGTAAGGCGGCTCGGAGAGTTACAAGACTGTAGAAACAACGGACTC
        Q  S  L  E  P  H  S  A  E  P  L  N  V  L  T  S  L  L  P  E

        CCTTGTTCGTGATGGTCGTATTGCTATTGTCAATGATTAACGTACTCCATCATGGTGGTG
   2582 --------+---------+---------+---------+---------+---------+-
2641
        GGAACAAGCACTACCAGCATAACGATAACAGTTACTAATTGCATGAGGTAGTACCACCAC
        P  C  S  *  W  S  Y  C  Y  C  Q  *  L  T  Y  S  I  M  V  V
```

```
        TTTCGAAGCACTCGTACCCATCTACCAACGCTAGTGGCAAACCTGTTCAAGGCACGTTAC
   2642 --------+---------+---------+---------+---------+---------+-
2701
        AAAGCTTCGTGAGCATGGGTAGATGGTTGCGATCACCGTTTGGACAAGTTCCGTGCAATG
        F  R  S  T  R  T  H  L  P  T  L  V  A  N  L  F  K  A  R  Y

        CAACCAAGCCGTGCCATCATCATCGTACTGTGGCTGTTGAGAAATGCTCCAACGACAACA
   2702 --------+---------+---------+---------+---------+---------+-
2761
        GTTGGTTCGGCACGGTAGTAGTAGCATGACACCGACAACTCTTTACGAGGTTGCTGTTGT
        Q  P  S  R  A  I  I  I  V  L  W  L  L  R  N  A  P  T  T  T

        CGGTACACTAGCCGGCATACCATGTCCTGTAAGCTTCTCCTACCTCTTCGGGGATGTCAC
   2762 --------+---------+---------+---------+---------+---------+-
2821
        GCCATGTGATCGGCCGTATGGTACAGGACATTCGAAGAGGATGGAGAAGCCCCTACAGTG
        R  Y  T  S  R  H  T  M  S  C  K  L  L  L  P  L  R  G  C  H

        CGAACATGATCTTGGTAGACCAGTGTAATAGCGCATAAAATCGCAATATCGCAACACAAG
   2822 --------+---------+---------+---------+---------+---------+-
2881
        GCTTGTACTAGAACCATCTGGTCACATTATCGCGTATTTTAGCGTTATAGCGTTGTGTTC
        R  T  *  S  W  *  T  S  V  I  A  H  K  I  A  I  S  Q  H  K

        CTTTTTCCATATACGGATCC
   2882 --------+---------+-  2901
        GAAAAAGGTATATGCCTAGG
        L  F  P  Y  T  D
```

Appendix 1.2 λBMKSA Sequence.

1.2.1 R156 (partial *pks* gene).

(Primers underlined)

```
        GAATTCTCTGGACCCAGCTtTACCAATGACACGGCCTGCTCCAGTTCGCTCGCAGCGATC
      1 --------+---------+---------+---------+---------+---------+
60
        CTTAAGAGACCTGGGTCGAaATGGTTACTGTGCCGGACGAGGTCAAGCGAGCGTCGCTAG
        E  F  S  G  P  S  F  T  N  D  T  A  C  S  S  S  L  A  A  I

        CATTTGGCTTGCAACTCGCTCTGGCGCGGCGATTGCGATACTGCTGTGGCGGGTGGCACG
     61 --------+---------+---------+---------+---------+---------+
120
        GTAAACCGAACGTTGAGCGAGACCGCGCCGCTAACGCTATGACGACACCGCCCACCGTGC
        H  L  A  C  N  S  L  W  R  G  D  C  D  T  A  V  A  G  G  T
```

```
      AACATGATCTTCACACCTGATGGTCACGCTGGTCTCGACAAAGGGTTCTTCCTGTCCCGT
  121 ---------+---------+---------+---------+---------+---------+
180
      TTGTACTAGAAGTGTGGACTACCAGTGCGACCAGAGCTGTTTCCCAAGAAGGACAGGGCA
      N  M  I  F  T  P  D  G  H  A  G  L  D  K  G  F  F  L  S  R


      ACTGGTAACTGTAAGCCTTTCGATGACAAGGCTGACGGATACTGTCGTGCTGAGGGTGTT
  181 ---------+---------+---------+---------+---------+---------+
240
      TGACCATTGACATTCGGAAAGCTACTGTTCCGACTGCCTATGACAGCACGACTCCCACAA
      T  G  N  C  K  P  F  D  D  K  A  D  G  Y  C  R  A  E  G  V


      GGTACCGTTATGGTCAAGAGGCTCGAAGATGCTCTTGCGGACGGAGATCCAATCCTTGGC
  241 ---------+---------+---------+---------+---------+---------+
300
      CCATGGCAATACCAGTTCTCCGAGCTTCTACGAGAACGCCTGCCTCTAGGTTAGGAACCG
      G  T  V  M  V  K  R  L  E  D  A  L  A  D  G  D  P  I  L  G


      ACGATCCTCGACGCGAAGACGAACCACTCCGCCATGAGCGACTCTATGACTCGCCCCTTC
  301 ---------+---------+---------+---------+---------+---------+
360
      TGCTAGGAGCTGCGCTTCTGCTTGGTGAGGCGGTACTCGCTGAGATACTGAGCGGGGAAG
      T  I  L  D  A  K  T  N  H  S  A  M  S  D  S  M  T  R  P  F


      GTCCCAGCCCAGATCGACAACATGGAAGCTTGCCTCAGCACCGCTGGAGTGGACCCTACC
  361 ---------+---------+---------+---------+---------+---------+
420
      CAGGGTCGGGTCTAGCTGTTGTACCTTCGAACGGAGTCGTGGCGACCTCACCTGGGATGG
      V  P  A  Q  I  D  N  M  E  A  C  L  S  T  A  G  V  D  P  T


      TCTCTCGACTACATTGAGATGCACGGTACTGGTACTCAAGTCGGCGACGCAGTCGAGATG
  421 ---------+---------+---------+---------+---------+---------+
480
      AGAGAGCTGATGTAACTCTACGTGCCATGACCATGAGTTCAGCCGCTGCGTCAGCTCTAC
      S  L  D  Y  I  E  M  H  G  T  G  T  Q  V  G  D  A  V  E  M


      GAGTCTGTTCTCAGCGTCTTTGCGCCGAATGAGCAGTTCCGCGGCAAGGACCAGCCTCTG
  481 ---------+---------+---------+---------+---------+---------+
540
      CTCAGACAAGAGTCGCAGAAACGCGGCTTACTCGTCAAGGCGCCGTTCCTGGTCGGAGAC
      E  S  V  L  S  V  F  A  P  N  E  Q  F  R  G  K  D  Q  P  L


      TATGTCGGCTCCGCCAAGGCCAACATCGGACACGGTGAGGGTGTGTCTGGTGTCACCAGT
  541 ---------+---------+---------+---------+---------+---------+
600
      ATACAGCCGAGGCGGTTCCGGTTGTAGCCTGTGCCACTCCCACACAGACCACAGTGGTCA
      Y  V  G  S  A  K  A  N  I  G  H  G  E  G  V  S  G  V  T  S
```

```
      TTGATCAAGGTCCTTCTCATGATGCAGACCAACCACTATCCGCCGCATTGCGGTATCAAG
  601 ---------+---------+---------+---------+---------+---------+
660                    SLpks4
      AACTAGTTCCAGGAAGAGTACTACGTCTGGTTGGTGATAGGCGGCGTAACGCCATAGTTC
      L  I  K  V  L  L  M  M  Q  T  N  H  Y  P  P  H  C  G  I  K

      CCTGGAAGCAAGATCAACCACAATTACCCGGATCTTGCGGCAAGAAATGTGCACATCGCG
  661 ---------+---------+---------+---------+---------+---------+
720
      GGACCTTCGTTCTAGTTGGTGTTAATGGGCCTAGAACGCCGTTCTTTACACGTGTAGCGC
      P  G  S  K  I  N  H  N  Y  P  D  L  A  A  R  N  V  H  I  A

      TTTGAGCCGAAACCGTTCTTGAGACGGGAGGGCAAGTTAAGACGGGTTTTGATCAATAAC
  721 ---------+---------+---------+---------+---------+---------+
780
      AAACTCGGCTTTGGCAAGAACTCTGCCCTCCCGTTCAATTCTGCCCAAAACTAGTTATTG
      F  E  P  K  P  F  L  R  R  E  G  K  L  R  R  V  L  I  N  N

      TTCAGTGCTGCAGGTGGCAATACTGCGCTTCTCATTGAGGATGCGCCTGACAGGATGCCG
  781 ---------+---------+---------+---------+---------+---------+
840
      AAGTCACGACGTCCACCGTTATGACGCGAAGAGTAACTCCTACGCGGACTGTCCTACGGC
      F  S  A  A  G  G  N  T  A  L  L  I  E  D  A  P  D  R  M  P

                                                          SlPKS2
      CTCTCAGGACAAGATCCTCGCACGACTCAGACTGTCACGATCTCGGGACATGTTGGCAAG
  841 ---------+---------+---------+---------+---------+---------+
900
      GAGAGTCCTGTTCTAGGAGCGTGCTGAGTCTGACAGTGCTAGAGCCCTGTACAACCGTTC
      L  S  G  Q  D  P  R  T  T  Q  T  V  T  I  S  G  H  V  G  K

      TCTCTCAGCAACAATGTCGCCAACTTGCTCGCACATCTGAAGAAGAATCCTACCATCGAT
  901 ---------+---------+---------+---------+---------+---------+
960
      AGAGAGTCGTTGTTACAGCGGTTGAACGAGCGTGTAGACTTCTTCTTAGGATGGTAGCTA
      S  L  S  N  N  V  A  N  L  L  A  H  L  K  K  N  P  T  I  D

      CTCTCACAGCTCGCCTACACGGTCAGTGCACGAAGATGGCATCACCTCCATCGTGTTGCT
  961 ---------+---------+---------+---------+---------+---------+
1020                                          slPKS1
      GAGAGTGTCGAGCGGATGTGCCAGTCACGTGCTTCTACCGTAGTGGAGGTAGCACAACGA
      L  S  Q  L  A  Y  T  V  S  A  R  R  W  H  H  L  H  R  V  A

      GTCGCGGGTACTACCGTCGCAGATATTACCGCGAAGTTGGAGAAAGCCATTGAGAATAAG
 1021 ---------+---------+---------+---------+---------+---------+
1080
      CAGCGCCCATGATGGCAGCGTCTATAATGGCGCTTCAACCTCTTTCGGTAACTCTTATTC
      V  A  G  T  T  V  A  D  I  T  A  K  L  E  K  A  I  E  N  K
```

```
        GAAGGTGTCAACAGACCTAAGGCGAAGCCTTCGGTCTTCTTCGCCTTCACAGGTCAAGGA
   1081 ---------+---------+---------+---------+---------+---------+
1140
        CTTCCACAGTTGTCTGGATTCCGCTTCGGAAGCCAGAAGAAGCGGAAGTGTCCAGTTCCT
        E  G  V  N  R  P  K  A  K  P  S  V  F  F  A  F  T  G  Q  G

        TCTCAGTACCTCGGCATGGGCAAGCAACTCTACGACTCTTATCCAATGTTCAGATCCGAG
   1141 ---------+---------+---------+---------+---------+---------+
1200
        AGAGTCATGGAGCCGTACCCGTTCGTTGAGATGCTGAGAATAGGTTACAAGTCTAGGCTC
        S  Q  Y  L  G  M  G  K  Q  L  Y  D  S  Y  P  M  F  R  S  E

        CTTCAAGGCTACGATCGCTTGGCACAATCGCAAGGCTTCCCAAGCTTTGCACACATCTTC
   1201 ---------+---------+---------+---------+---------+---------+
1260
        GAAGTTCCGATGCTAGCGAACCGTGTTAGCGTTCCGAAGGGTTCGAAACGTGTGTAGAAG
        L  Q  G  Y  D  R  L  A  Q  S  Q  G  F  P  S  F  A  H  I  F

                                                                SLpks3
        ACCGAGACGAAGGGAGATGTTGAACAGAATCTTCCAGTGGTCGTGCAGCTTGCTATTACA
   1261 ---------+---------+---------+---------+---------+---------+
1320
        TGGCTCTGCTTCCCTCTACAACTTGTCTTAGAAGGTCACCAGCACGTCGAACGATAATGT
        T  E  T  K  G  D  V  E  Q  N  L  P  V  V  V  Q  L  A  I  T

        TGCTTGCAAATGGCTCTCTTCAACCTCGTCACCTCCTTCGGAATCAAGGCCTCTGCCGTT
   1321 ---------+---------+---------+---------+---------+---------+
1380
        ACGAACGTTTACCGAGAGAAGTTGGAGCAGTGGAGGAAGCCTTAGTTCCGGAGACGGCAA
        C  L  Q  M  A  L  F  N  L  V  T  S  F  G  I  K  A  S  A  V

        GTCGGCCACTCGCTGGGCGAGTACGCTGCGCTGTATGCAGCTGGTGTGTTGAGTGCCAGC
   1381 ---------+---------+---------+---------+---------+---------+
1440
        CAGCCGGTGAGCGACCCGCTCATGCGACGCGACATACGTCGACCACACAACTCACGGTCG
        V  G  H  S  L  G  E  Y  A  A  L  Y  A  A  G  V  L  S  A  S

        GACACGATCTACCTGGTCGGCAAACGTGCCGAGCTTCTCCAGGATCATTGCCAGAGGGGT
   1441 ---------+---------+---------+---------+---------+---------+
1500
        CTGTGCTAGATGGACCAGCCGTTTGCACGGCTCGAAGAGGTCCTAGTAACGGTCTCCCCA
        D  T  I  Y  L  V  G  K  R  A  E  L  L  Q  D  H  C  Q  R  G

        ACGCATGCGATGCTTGCGTGCAAGGCGAGTGAGTGGAGTCTCGCCGAGATCACGGCGGGC
   1501 ---------+---------+---------+---------+---------+---------+
1560
        TGCGTACGCTACGAACGCACGTTCCGCTCACTCACCTCAGAGCGGCTCTAGTGCCGCCCG
        T  H  A  M  L  A  C  K  A  S  E  W  S  L  A  E  I  T  A  G
```

```
        AAGAATGTCGAAGTCGCATGCGTTAATGGGCCTGAAGACACTGTCCTCTCCGGCACTGTC
   1561 ---------+---------+---------+---------+---------+---------+
1620
        TTCTTACAGCTTCAGCGTACGCAATTACCCGGACTTCTGTGACAGGAGAGGCCGTGACAG
        K  N  V  E  V  A  C  V  N  G  P  E  D  T  V  L  S  G  T  V

        GAGGAAATTGGAGAGGTGCAGAAGACACTCATTGCGAAGAGCATCAAGGCTACACTCTTG
   1621 ---------+---------+---------+---------+---------+---------+
1680
        CTCCTTTAACCTCTCCACGTCTTCTGTGAGTAACGCTTCTCGTAGTTCCGATGTGAGAAC
        E  E  I  G  E  V  Q  K  T  L  I  A  K  S  I  K  A  T  L  L

        AAGTTGCCCTTCGCGTTTCATTCGGCGCAGGTACAACCTATCCTCCGAGGACGTTCTGAA
   1681 ---------+---------+---------+---------+---------+---------+
1740
        TTCAACGGGAAGCGCAAAGTAAGCCGCGTCCATGTTGGATAGGAGGCTCCTGCAAGACTT
        K  L  P  F  A  F  H  S  A  Q  V  Q  P  I  L  R  G  R  S  E

        GAACTTGCGGCTGGAGCTACTTTTGAGAAGCCCAAGCTTGCGGTCATTTCCCCGCTACTG
   1741 ---------+---------+---------+---------+---------+---------+
1800
        CTTGAACGCCGACCTCGATGAAAACTCTTCGGGTTCGAACGCCAGTAAAGGGGCGATGAC
        E  L  A  A  G  A  T  F  E  K  P  K  L  A  V  I  S  P  L  L

        GGCAGTGTGGTCGACGACGAAGGAGTCGTTGGACCCAACTACCTTGCACGCCACTGCCGT
   1801 ---------+---------+---------+---------+---------+---------+
1860
        CCGTCACACCAGCTGCTGCTTCCTCAGCAACCTGGGTTGATGGAACGTGCGGTGACGGCA
        G  S  V  V  D  D  E  G  V  V  G  P  N  Y  L  A  R  H  C  R

        GAGGCGGTCGGAATGGTCAAAGCCCTCGGAGTGGCGAAGGAGAAGGGTATAATCAACGAG
   1861 ---------+---------+---------+---------+---------+---------+
1920
        CTCCGCCAGCCTTACCAGTTTCGGGAGCCTCACCGCTTCCTCTTCCCATATTAGTTGCTC
        E  A  V  G  M  V  K  A  L  G  V  A  K  E  K  G  I  I  N  E

        AAGACCTTCGTCATTGAGATTGGTCCTAAGCCGCTTCTCTGCGGAATGATCAAGAACATA
   1921 ---------+---------+---------+---------+---------+---------+
1980
        TTCTGGAAGCAGTAACTCTAACCAGGATTCGGCGAAGAGACGCCTTACTAGTTCTTGTAT
        K  T  F  V  I  E  I  G  P  K  P  L  L  C  G  M  I  K  N  I

        CTCGGCCAGAACATCGTAGCCTTGCCTACGTTGAAGGACAAGGGTCCAGACGTCTGGCAG
   1981 ---------+---------+---------+---------+---------+---------+
2040
        GAGCCGGTCTTGTAGCATCGGAACGGATGCAACTTCCTGTTCCCAGGTCTGCAGACCGTC
        L  G  Q  N  I  V  A  L  P  T  L  K  D  K  G  P  D  V  W  Q
```

```
        AACCTCTCGAACATCTTCACGACGCTCTACACCGGTGGTTTAGACATCAACTGGACTGCC
   2041 ---------+---------+---------+---------+---------+---------+
2100
        TTGGAGAGCTTGTAGAAGTGCTGCGAGATGTGGCCACCAAATCTGTAGTTGACCTGACGG
        N  L  S  N  I  F  T  T  L  Y  T  G  G  L  D  I  N  W  T  A

        TTCCACGCCCCCTTCGAGCCCGCGAAGAAGGTCCTGCAACTTCCTGATTATGGCTGGGAT
   2101 ---------+---------+---------+---------+---------+---------+
2160
        AAGGTGCGGGGGAAGCTCGGGCGCTTCTTCCAGGACGTTGAAGGACTAATACCGACCCTA
        F  H  A  P  F  E  P  A  K  K  V  L  Q  L  P  D  Y  G  W  D

        CTCAAGGATTACTTCATCCAGTATGAAGGCGATTGGGTTCTGCATCGGCACAAGATCCAC
   2161 ---------+---------+---------+---------+---------+---------+
2220
        GAGTTCCTAATGAAGTAGGTCATACTTCCGCTAACCCAAGACGTAGCCGTGTTCTAGGTG
        L  K  D  Y  F  I  Q  Y  E  G  D  W  V  L  H  R  H  K  I  H

        TGCAACTGTGCAGATGCTGGAAAGGATGTGCATAACACTTCGCACTACTGTCCTGGCAAA
   2221 ---------+---------+---------+---------+---------+---------+
2280
        ACGTTGACACGTCTACGACCTTTCCTACACGTATTGTGAAGCGTGATGACAGGACCGTTT
        C  N  C  A  D  A  G  K  D  V  H  N  T  S  H  Y  C  P  G  K

        CACACCTTCGCTGAGAATGTTGTCGTTCCTGGTGGGGCTCAGAAGGCCGTTCAGGAAGCA
   2281 ---------+---------+---------+---------+---------+---------+
2340
        GTGTGGAAGCGACTCTTACAACAGCAAGGACCACCCCGAGTCTTCCGGCAAGTCCTTCGT
        H  T  F  A  E  N  V  V  V  P  G  G  A  Q  K  A  V  Q  E  A

        CCTGCGGCGAAGACAGAGACGAAGAAGATGTCGAAGCTGGATCC
   2341 ---------+---------+---------+---------+----  2384
        GGACGCCGCTTCTGTCTCTGCTTCTTCTACAGCTTCGACCTAGG
        P  A  A  K  T  E  T  K  K  M  S  K  L  D
```

1.2.2 R161 Sequence generated with pUCfwd primer.

1.2.3 R162 Sequence generated with pUCfwd primer.

```
1   GCAGGTCGAC TCTAGAGGAT CCTTGACGTG GCACAGGCAC TGGCACTGGC

51  ACTGGCACTG GCACTGGCAC TGGCACTGGC ACAGACACTG GCACTCGCAC

101 TGGCACTGGG CTCGCACTGC TCTGAACGAC CTTGACCTCG CACCCGTGAA

151 CGAACAGCTT CATCACAGCT CATCGTTCAT AATAACCAGA TCAAGCCATC

201 AATATGCCCA CCTATGCCCT GCTCGGCGCG ACAGGCGCCA CAGGCTCGGC

251 CATACTCCGC TGCCTCCTTG CCTCACCACC CCCAGACCTC GACCTCAACA
```

301 TCCTCGTCCG CTCCAAACAA AAGCTCCTCA AGTCCTTCCC AACCCTCACT

351 ACAACAATCT CTCCCCGCAT CCACATCATC CAAGGCAACT CCACCGACAC

401 CATCGCTCTT CAACAATGTC TCGAAACGCG TCCGTAGCCT TCATGTGCGT

451 CGCCGACAAT GCCTCCAACA AAGGGTGTCT CTCTACC

1.2.4 R162 Sequence generated with reverse primer.

1 ACGAATTCGA GCTCGGTACC CGGGGATCC AGCATAGAAT CCACCCACAT

51 GCCGTTCACC TGCTCTTCGA TCTCTCTGGC GTGAGCAACG ACGGCTTTGG

101 GATCGCGACG GAAGTTCTCT CGCTCTTCTT CGCTGTACTC TTTGGCTTGA

151 CTACCAGAGT TTCCGGCTAT CACCCCGAAC CAGACTCCCG TACGGACGAA

201 GATGTCGAGG TGTTTTGCGT ATGGCTGCAT GCCTGGAACA GTCTGTATGG

251 ATGAAGCACC AGATCCAATT ACGGCGACTC GCTCCTTTGC CCAGTCTTCC

301 TCCTTGTAGT CCTTTGGCCA CCGGGCGGTG TGAATGACCC GGCCCTTGAA

351 GCGATCCTGC AGGCCTGGAA TATCGGGGAA CTGGTGGCTC ACGTTAGCAC

401 AAATCTCGAC TCCGGACTGT GTCAAACATA CCTTGAAATT ATTCAAGCAC

451 GCCAGCGCCA TACAACAGCA CATTACATCT ATCCTCGAAC TCCCGGCCTC

501 GGGGGTTGGA AAAATGGTCT CGAAGCTTGA CACCCATTCG CCGCATGCTC

551 TGGCCAGTAC ACCCCACAAC TGCACGTGGA CC

1.2.5 R163 sequence generated using pUCfwd primer.

1 CCTGCAGGTC GACTCTAGAG GATCCCCGGG TACCCGAGCT CGAATTCTCC

51 GGCGAAATCG AGCCAGCAGC CCCATTTGGT CGCACCCTTG TTGCGAGCAC

101 GCCCAGTTGG GTCGACGTGA GTCTCGACGT CCCAACGACG TTTGGGCACT

151 TCCTTGACCA CGTCAAGCCC CTTGTAGAGC AGATCCCAAA AAGCGTCGGT

201 GCTTTGTGCC TCTGGGAAGC GACCAGACAT TGAGACGATA GCAAGCTTGC

251 TCTTTCCGAA CCAGGTCTAT GGGCGAAAGA TTCGTTGATT GTGGGCTTCT

301 CGACCTCGAT ATTGGAGATC GTTGGGCTGA GTGCAGCCGA CAGAGCACGG

351 TCGGCAGAGG TGGTGATTGG GCGTGATGAG CACATTTTCT GCGCCGCGAG

401 ACTTGAGGAG CCTTGGGAGC TCAANACTCG ACCTTACCCC AGCCAATGCT

451 CTCGAGCAGG CACTGATTGA GGGCGAGATG GATTACTGCA G

Appendix 2: Codon Usage Tables.

2.1 Codon Usage in *D. pini phn*1.

A.A	Codon	Number	/1000	Fraction	A.A	Codon	Number	/1000	Fraction
Gly	GGG	1.00	3.77	0.50	Trp	TGG	6.00	22.64	1.00
Gly	GGA	0.00	0.00	0.00	End	TGA	3.00	11.32	0.75
Gly	GGT	1.00	3.77	0.50	Cys	TGT	4.00	15.09	0.67
Gly	GGC	0.00	0.00	0.00	Cys	TGC	2.00	7.55	0.33
Glu	GAG	2.00	7.55	1.00	End	TAG	0.00	0.00	0.00
Glu	GAA	0.00	0.00	0.00	End	TAA	1.00	3.77	0.25
Asp	GAT	2.00	7.55	1.00	Tyr	TAT	1.00	3.77	1.00
Asp	GAC	0.00	0.00	0.00	Tyr	TAC	0.00	0.00	0.00
Val	GTG	13.00	49.06	0.72	Leu	TTG	14.00	52.83	0.42
Val	GTA	2.00	7.55	0.11	Leu	TTA	2.00	7.55	0.06
Val	GTT	1.00	3.77	0.06	Phe	TTT	0.00	0.00	0.00
Val	GTC	2.00	7.55	0.11	Phe	TTC	2.00	7.55	1.00
Ala	GCG	10.00	37.74	0.34	Ser	TCG	19.00	71.70	0.38
Ala	GCA	8.00	30.19	0.28	Ser	TCA	12.00	45.28	0.24
Ala	GCT	3.00	11.32	0.10	Ser	TCT	5.00	18.87	0.10
Ala	GCC	8.00	30.19	0.28	Ser	TCC	3.00	11.32	0.06
Arg	AGG	13.00	49.06	0.48	Arg	CGG	0.00	0.00	0.00
Arg	AGA	11.00	41.51	0.41	Arg	CGA	1.00	3.77	0.04
Ser	AGT	2.00	7.55	0.04	Arg	CGT	2.00	7.55	0.07
Ser	AGC	9.00	33.96	0.18	Arg	CGC	0.00	0.00	0.00
Lys	AAG	3.00	11.32	0.75	Gln	CAG	1.00	3.77	0.14
Lys	AAA	1.00	3.77	0.25	Gln	CAA	6.00	22.64	0.86
Asn	AAT	2.00	7.55	1.00	His	CAT	1.00	3.77	1.00
Asn	AAC	0.00	0.00	0.00	His	CAC	0.00	0.00	0.00
Met	ATG	4.00	15.09	1.00	Leu	CTG	11.00	41.51	0.33
Ile	ATA	0.00	0.00	0.00	Leu	CTA	1.00	3.77	0.03
Ile	ATT	0.00	0.00	0.00	Leu	CTT	3.00	11.32	0.09
Ile	ATC	1.00	3.77	1.00	Leu	CTC	2.00	7.55	0.06
Thr	ACG	12.00	45.28	0.36	Pro	CCG	8.00	30.19	0.26
Thr	ACA	7.00	26.42	0.21	Pro	CCA	11.00	41.51	0.35
Thr	ACT	9.00	33.96	0.27	Pro	CCT	3.00	11.32	0.10
Thr	ACC	5.00	18.87	0.15	Pro	CCC	9.00	33.96	0.29

2.2 Codon Usage in Partial *D. pini pks* Sequence.

AmAcid	Codon	Number	/1000	Fraction
Gly	GGG	3.00	3.78	0.04
Gly	GGA	20.00	25.19	0.29
Gly	GGT	24.00	30.23	0.34
Gly	GGC	23.00	28.97	0.33
Glu	GAG	27.00	34.01	0.68
Glu	GAA	13.00	16.37	0.32
Asp	GAT	21.00	26.45	0.50
Asp	GAC	21.00	26.45	0.50
Val	GTG	11.00	13.85	0.19
Val	GTA	2.00	2.52	0.03
Val	GTT	14.00	17.63	0.24
Val	GTC	32.00	40.30	0.54
Ala	GCG	27.00	34.01	0.34
Ala	GCA	14.00	17.63	0.17
Ala	GCT	20.00	25.19	0.25
Ala	GCC	19.00	23.93	0.24
Arg	AGG	3.00	3.78	0.12
Arg	AGA	6.00	7.56	0.23
Ser	AGT	8.00	10.08	0.17
Ser	AGC	9.00	11.34	0.19
Lys	AAG	45.00	56.68	0.88
Lys	AAA	6.00	7.56	0.12
Asn	AAT	13.00	16.37	0.37
Asn	AAC	22.00	27.71	0.63
Met	ATG	17.00	21.41	1.00
Ile	ATA	2.00	2.52	0.05
Ile	ATT	10.00	12.59	0.26
Ile	ATC	26.00	32.75	0.68
Thr	ACG	15.00	18.89	0.31
Thr	ACA	6.00	7.56	0.12
Thr	ACT	14.00	17.63	0.29
Thr	ACC	14.00	17.63	0.29
Trp	TGG	7.00	8.82	1.00

AmAcid	Codon	Number	/1000	Fraction
End	TGA	0.00	0.00	0.00
Cys	TGT	4.00	5.04	0.25
Cys	TGC	12.00	15.11	0.75
End	TAG	0.00	0.00	0.00
End	TAA	0.00	0.00	0.00
Tyr	TAT	6.00	7.56	0.32
Tyr	TAC	13.00	16.37	0.68
Leu	TTG	13.00	16.37	0.18
Leu	TTA	2.00	2.52	0.03
Phe	TTT	6.00	7.56	0.20
Phe	TTC	24.00	30.23	0.80
Ser	TCG	10.00	12.59	0.21
Ser	TCA	2.00	2.52	0.04
Ser	TCT	10.00	12.59	0.21
Ser	TCC	8.00	10.08	0.17
Arg	CGG	3.00	3.78	0.12
Arg	CGA	2.00	2.52	0.08
Arg	CGT	6.00	7.56	0.23
Arg	CGC	6.00	7.56	0.23
Gln	CAG	18.00	22.67	0.64
Gln	CAA	10.00	12.59	0.36
His	CAT	11.00	13.85	0.41
His	CAC	16.00	20.15	0.59
Leu	CTG	10.00	12.59	0.14
Leu	CTA	1.00	1.26	0.01
Leu	CTT	15.00	18.89	0.21
Leu	CTC	32.00	40.30	0.44
Pro	CCG	9.00	11.34	0.22
Pro	CCA	6.00	7.56	0.15
Pro	CCT	18.00	22.67	0.45

References.

Al-Samarrai, T., Schmid, J. (1999) A simple method for the extraction of fungal genomic DNA. *Letters in Applied Microbiology.* (In press).

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

Balance, D.J. (1986). Sequences important for gene expression in filamentous fungi. *Yeast.* **2**, 229-236.

Bhatnagar, D., Ehrlich, K.C., Cleveland, T.E. (1992) Oxidation-reduction reactions in biosynthesis of secondary metabolites and their control by environmental factors. *Handbook of Applied Microbiology.* **V**, 255-286, Marcel Dekker Inc. New York.

Brown, D. W., Yu, J.-H., 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.

Bullock, W.O., Fernandez, J.M., and Short, J.M. (1987). XL-1-Blue: a high efficiency plasmid transrorming *rec*A *Escherichia coli* strain with β-galacyosidase selection. *Biotechniques.* **5**, 376-378.

Bulman, L. (1992). Forestry Quarantine Risk of Cargo Imported into New Zealand. *New Zealand Journal of Forestry Science.*
22(1), 32-38.

Carson, S. D. (1989). Selecting Pinus radiata for resistance to Dothistroma needle blight. *New Zealand Journal of Forestry Science* **19**(1), 3-21.

Carson, S. D. (1991). Genotype X environment interaction and optimal number of progeny test sites for improving Pinus radiata in New Zealand. *New Zealand Journal of Forestry Science* **21**(1), 32-49.

Chang, P. K., Skory, C. D. and Linz, J. E. (1992). Cloning of a gene associated with aflatoxin B1 biosynthesis in *Aspergillus parasiticus. Current Genetics* **21**, 231-233.

Chang, P. K., 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**(10), 3273-9.

Chang, P. K., Ehrlich, K. C., 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.

Chang, P. K., Cary, J. W., Yu, J., Bhatnagar, D. and Cleveland, T. E. (1995). The Aspergillus parasiticus polyketide synthase gene pksA, a homolog of Aspergillus nidulans wA, is required for aflatoxin B1 biosynthesis. *Molecular General Genetics* **248**(3), 270-277.

Chou, C. (1991). Perspectives of disease threat in Large Scale *Pinus radiata* Monoculture- The New Zealand Experience. *European Journal of Forest Pathology.* **21**, 71-81.

Cleveland, T. E., Bhatnagar, D. and Brown, R. L. (1991). Aflatoxin production via cross-feeding of pathway intermediates during cofermentation of aflatoxin pathway-blocked Aspergillus parasiticus mutants. *Applied and Environmental Microbiology* **57**(10), 2907-2911.

Cohen, S.N., Chang, A.C.Y., Hsu, L. (1972). Nonchromosomal antibiotic resistance in bacteria: Genetic transformation of *E. coli* by R-factor DNA. *Proceedings of The National Academy of Sciences.* **69**, 2110.

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**(2/3), 171-179.

Drew, S., Demain, A.(1977) Effect of Primary Metabolites on Secondary Metabolism. *Annual Review of Microbiology.*
31,343-356.

Edwards, M., Bowling, D.(1986). The Growth of Rust Germ Tubes Toward Stomata in Relation to pH Gradients. *Physiological and Molecular Plant Pathology.*
29, 185-196.

Feng, G.H., Leonard, T.J. (1995) Characterisation of the polyketide synthase gene (*pks*-1) required for aflatoxin biosynthesis in *Aspergillus parasiticus.* *Journal of Bacteriology.* **177**, 6246-6254.

Ferguson, L. R., Parslow, M. I. and McLarin, J. A. (1986). Chromosome damage by dothistromin in human peripheral blood lymphocyte cultures: a comparison with aflatoxin B1. *Mutat Res* **170**(1-2), 47-53.

Forestry Research Institute. Control Of *Dothistroma pini.*
F.R.I. Bulletin number 18.

Franich, R., Wells, L., Barnett, J.(1977). Variation With Tree Age of Needle Cuticle Topography and Stomatal Structure in *Pinus radiata* D.Don. *Annual of Botany.*
41, 621-626.

Franich, R., Gaskin, L., Wells, L., Zabekiewicz, J.(1982). Effect of *Pinus radiata* needle Monoterpenes on Spore Germination and Mycelial Growth of *Dothistroma pini in vitro* in relation to Mature Tree Resistance. *Physiological Plant Pathology.*
21, 55-63.

Franich, R.A., Gadgil, P.D., Shain, L. (1983). Fungistatic Effects of *Pinus radiata* Needle Epiticular Fatty and Resin acids on *Dothistroma pini*. *Physiological Plant Pathology.*
23, 183-195.

Franich, R. J., Carson, M. J. and Carson, S. D. (1986). Synthesis and accumultation of benzoic acid in response to tissue injury by dothistromin, and correlation with resistance of *Pinus radiata* families to *Dothistroma pini*. *Physiological and Molecular Plant Pathology* **28**(2), 267-286.

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**(3), 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.(1969). Survival of *Dothistroma pini* on Fallen Needles of *Pinus radiata*. *New Zealand Journal of Botany.*
8,303-309.

Gadgil, P.(1974). Effect of Temperature and Leaf Wetness period on Infection of *Pinus radiata* by *Dothistroma pini*. *New Zealand Journal of Forestry Science.*
4(3), 495-501.

Gadgil, P. D. and Holden, G. (1976). Effect of light intensity on infection of Pinus radiata by Dothistroma pini. *New Zealand Journal of Forestry Science* **6**(1), 67-71.

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

Gallagher, R.T., Hodges, R. (1972) The chemistry of dothistromin, a difuroanthraquinone from *Dothistroma pini*. *Austalian Journal of Chemistry*. **25**, 2399-2407.

Gibson, I. A. S. (1972). Dothistroma blight of *Pinus radiata*. *Annual Review of Phytopathology* **10**, 51-72.

Gillman, C.(1996). Identification of Dothistromin Biosynthetic Pathway Genes.
Msc. Thesis.
Massey University, Palmerston North.

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

Heath, M.C. (1996) Plant resistance to fungi. *Canadian Journal of Plant Pathology*. **18**, 469-475.

Hirst, P.(1997). Genetic Diversity of *Dothistroma pini* in New Zealand.
MSc. Thesis.
Massey University, Palmerston North.

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

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

Keller, N.P., Segner, S., Bhatnager, D., and Adams, T.H. (1995). *Stc*S, 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., Hohn, T.M. (1997). Metabolic Pathway Gene Clusters in Filamentous Fungi. *Fungal Genetics and Biology.*
21, 17-29.

Leder, P., Tiemeier, D., Enquist, L. (1977). EK2 derivatives of bacteriophage lambda useful in the cloning of DNA from higher organisms: the λgtWES system. *Science*. **196**, 175.

Liang, S. H., 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**(12), 4568-4575.

Lloyd, A.T., Sharp, P.M. (1991) Codon usage in *Aspergillus nidulans*. *Molecular Journal of Genetics.* **230**, 288-294.

Martin, J., Liras, P.(1989). Organisation and Expression of Genes Involved in the Biosynthesis of Antibiotics and Other Secondary Metabolites. *Annual Review of Microbiology.*
43, 173-206.

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

Ministry Of Forestry (1996). Forestry Facts and Figures

Moffat, A.(1996). Moving Forest Trees into The Genetics Era. *Science.*
271, 760-761.

Monahan.(1998). Identification of putative Dothistromin Biosynthetic Genes.
Msc. Thesis.
Massey University, Palmerston North.

Morgan.(1997). Identification of a polyketide synthase gene from *Dothistroma pini.*
Msc. *Thesis.*
Massey University, Palmerston North.

Murray, N.E.W., Brammer, W.J., Murray. (1977). Lambdoid phages that simplify the recovery of *in vitro* recombinants. *Molecular and General Genetics*. **150**, 53-58.

New, D. (1989). Forest health - an industry perspective of the risks to New Zealand's plantations. *Special issue: Workshop on Forest Health in the South Pacific, Rotorua* **19**, 2-3.

Proceedings of the International Union of Forest Research Organisations (IUFRO) (1997) Quebec City, May 25-31.

Sambrook, J., Fritch, E.F., Maniatis, T. (1989). *Molecular cloning:A laboratory manual.* New York: Cold Spring Harbour Laboratory press.

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

Shaw, C., Toes, E. (1977). Impact of Dothistroma Needle Blight and Armillaria root rot on Diameter Growth of Pinus Radiata. *Phytopathology*.
67, 1319-1323.

Sheridan, J. E. and Yen, C. C. (1970). A note on the effect of temperature and reletive humidity on the germination of conida of a New Zealand isolate of Dothistroma pini hulbary. *New Zealand Journal of Botany* **8**, 658-660.

Shimizu, K., Tanaka, C., Tsuda, M. (1997). Cloning of *Brn*1, a reductase gene involved in melanin biosynthesis in *Cochliobolus heterostrophus*. *Lournal of General and Applied Microbiology*. **43**, 145-150.

Silva, J.C., Minto, R.E., Barry III,C.E ., Holland, K.A.(1996). Isolation and characterisation of the Versicolorin B Synthase gene from *Aspergillus parasiticus*. *The Journal of Biological Chemistry.*
271(23),13600-13608.

Skory, C. D., Chang, P. K., Cary, J. and Linz, J. E. (1992). Isolation and characterization of a gene from Aspergillus parasiticus associated with the conversion of versicolorin A to sterigmatocystin in aflatoxin biosynthesis. *Applied and Environmenyal Microbiology* **58**(11), 3527-37.

Stoessl, A. (1984). Dothistromin as a metabolite of *Cercospora arachidicola*. *Mycopathologia* **112**, 179-186.

Takano, Y., Kubo, Y., Shimiizu, K., Mise, K., Okuno, T., Furusawa, I. (1995) Structural analysis of *PKS*1, a polyketide synthase gene involved in melanin biosynthesis in *Colletotrichum lagenarium*. *Molecular General Genetics*. **249**, 162-167.

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**(10), 2940-2945.

Vidal-Cross, A., Viviani, F., Labesse, G., Bocara, M., Gaudry, M. (1994) Polyhydroxynapthalene reductase involved in melanin biosynthesis in *Magnaporthe grisea. European Journal of Biochemistry.* **2/9**, 985-992.

Weigel, B., Burgett, S., Chen, V., Skatrud, P., Frolik, C., Queener, S., Ingolia, T.(1988). Cloning and expression in *Escherichia coli* of Isopenicillin N Synthase genes From *Streptomyces lipmznii* and *Aspergillus nidulans. Journal of Bacteriology.* **170,** 3817-3826

Wolfe, H., Shields, D.(1997). Molecular Evidence for an Ancient Duplication of the Entire Yeast Genome. *Nature.* **387,** 708-713.

Woolons, R., Hayward, J.(1984). Growth Losses in *Pinus radiata* Stands Unsprayed for *Dothistroma pini. New Zealand Journal of Forestry Science.* **14**(1), 14-22.

Yu, J., Chang, P., Cary, J., Wright, M., Bhatnager, D., Cleveland, T., Payne, G., Linz, J. (1995). Comparative Mapping of Aflatoxin Pathway Gene Clusters in *Aspergillus parasiticus* and *Aspergillus flavus. Applied and Environmental Microbiology.* **61**(6), 2365-2371.

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