# *Paxilline negative mutants of Penicillium paxilli generated by heterologous and homologous plasmid integration*

A Thesis presented in partial fulfilment of

the requirements for the degree of

Master of Science in Molecular Genetics

at Massey University, Palmerston North,

New Zealand.

*Carolyn Anne Young*

*1998*

*This Thesis is dedicated to*

*My Father*

*Tui Dene Young*

*1935 - 1992*

*Dad I wish you could be here today.*
*I know how proud you would be.*
*I miss you heaps and*
*Love you forever.*

*Carolyn*

# *Abstract*

Using a monoclonal antibody-based ELISA, 600 pAN7-1 plasmid-tagged mutants of *Penicillium paxilli* were screened for paxilline accumulation and one paxilline negative mutant, YI-20, was identified (Itoh, unpublished data). A molecular analysis of this mutant showed that pAN7-1 was inserted at a single site but was present as 4-6 copies arranged in a head-to tail tandem repeat. Rescue of flanking sequences and analysis of the corresponding genomic region revealed that YI-20 has an extensive deletion at the site of pAN7-1 integration. Probing of a CHEF gel with the same sequences showed that associated with the deletion is a rearrangement of chromosome Va. Targeted gene disruption of wild-type sequences adjacent to the site where pAN7-1 inserted, resulted in the generation of two additional paxilline-negative mutants; both were single crossovers with deletions extending outside the region mapped. Neither of these new mutants had a rearrangement of chromosome Va, suggesting that deletion of genes on this chromosome is responsible for the paxilline-negative phenotype.

# *Acknowledgments*

I would like to express my sincerest thanks to Barry Scott not only as supervisor for my MSc but also as my boss of 11 years. Barry without your experience, guidance and patience over the years, this project, and many others would not have been as enjoyable or rewarding. Thanks for believing in me, for giving me the freedom I required and letting me get my own way (most of the time....).

To Yasuo Itoh, who began this work, it was sad to see you leave New Zealand but thank you for YI-20 and leaving behind such an interesting and challenging project. To Richard Johnson, together we made a terrible team, but, I definitely could not have done this without you. Thank you Rich for being there in the beginning and especially all our late night sessions. To Mike Collett, Austen Ganley, Rich and Yasuo, one of the many Scott Bases of old, thanks for listening, thanks for your input and more importantly thanks for all the nights at the staff club.

To all those people who have influenced me in my career, my research and my life, especially, Trish McClenachan, James Dickson, Kathryn Stowell, John Tweedie, Tim Brown, Miranda Clayton, Paul Reynolds, Kevin Farnden and Eva Rockman, thanks. Thanks to all the members, too many to name here, of Scott Base and the MGU, old and new, who have put up with me over the years, especially Trish, Lisa, Jo, Linda, Carmel, Emily, Christina, Brendon, Bek, Rosie, Max, Grant and Dianne.

Thank you to Ian Garthwaite, Kath Ross, Sarah Munday-Finch and Chris Miles, from Ruakura, for the ELISA and HPLC analysis.

To Geoff Jameson, thank you for always listening to me while I ranted on and on about my work and a especial thank you for proof reading this thesis.

Thank you to those involved at the 11 th hour. David, how could I have ever thought I could put all this together with out your help. Emily, thanks for making sure Patrick wasn't a "home alone" case. Lisa, thanks for being "lab mother". Barry, thank you thank you, thank you.

Finally, and most importantly to David and Patrick McSweeney, without your love, support, encouragement and patience I could never have achieved this. All the late nights are now over (well, perhaps not all...).

# Table of Contents

# *List of Tables*

# List of Figures

# *Chapter One*
# *Introduction*

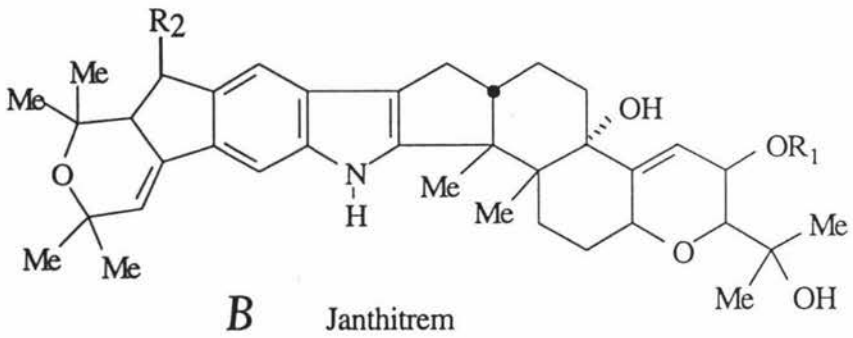
## 1.1 *Penicillium paxilli*

Members of the genus *Penicillium* are widespread throughout nature and most well known for their useful metabolite, the antibiotic penicillin. More recently, additional secondary metabolites have been isolated from *Penicillium*. Such metabolites have no known role in the organism's living processes (e.g. structure, replication, differentiation, communication and/or homeostasis) but are synthesized from either intermediates or end products of primary metabolism (Peberdy, 1987). A large group of these secondary metabolites, found not only in the *Penicillium spp.* but also in other filamentous fungi such as *Aspergillus spp.* (Gallagher and Wilson, 1978; Seya *et al.*, 1986; Nozawa *et al.*, 1987), *Claviceps spp.* (Cole *et al.*, 1977; Dorner *et al.*, 1984) and *Neotyphodium spp.* (Gallagher *et al.*, 1984) are known as tremorgenic mycotoxins. These toxins are fungal metabolites that affect the central nervous system function in vertebrates. An example of such a tremorgen is the indole-diterpenoid, paxilline, (Fig. 1.1) produced by *Penicillium paxilli*.

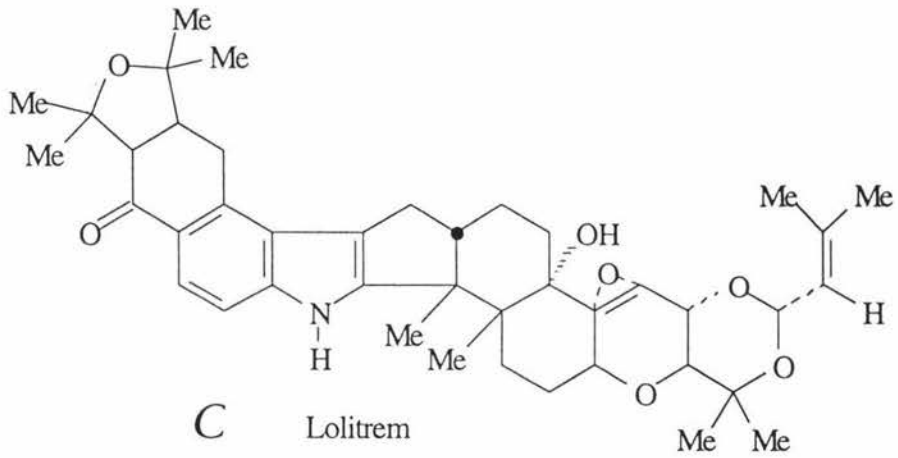
*Penicillium paxilli,* an asexual saprophytic fungi, was first isolated from insect-damaged pecans in the state of Georgia, USA (Cole *et al.*, 1974). This fungus grows rapidly, sporulates readily and produces large amounts of paxilline in submerged culture. Production of paxilline by *P. paxilli* was found to cause tremors when fed to both cockerels and mice (Cole *et al.*, 1974). Also reported, is the isolation of *P. paxilli* from the faeces of cattle in Victoria, Australia and soil from pastures in South Australia (Cockrum *et al.*, 1979). However, these two isolates did not produce the characteristic paxilline but substantial quantities of the tremorgenic toxin verruculogen, a prenylated diketopiperazine (Cockrum *et al.*, 1979). Although structurally different from paxilline, verruculogen contains both indole and isoprene moieties.

UV and IR spectra show that paxilline contains indole (Cole *et al.*, 1974) and its structure was determined by X-ray diffraction analysis. Data from labelling studies (Acklin *et al.*, 1977; de Jesus *et al.*, 1983; Laws and Mantle, 1989) show that both radiolabelled tryptophan and mevalonic acid are primary precursors of paxilline, where the incorporation of four mevalonate molecules indicates an involvement in isoprenoid biosynthesis (Fig. 1.2).

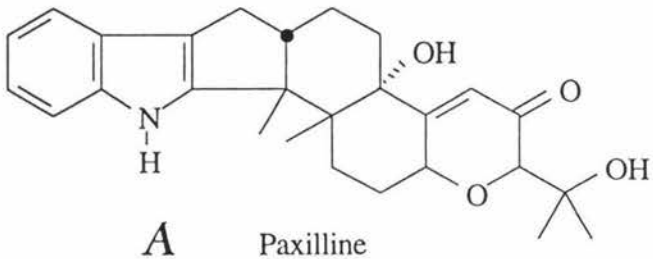
Secondary metabolites that contain tryptophan derivatives usually retain the aliphatic side chain. However, this tremorgenic mycotoxin has lost both the $\alpha$- and $\beta$-carbons (Peberdy, 1987). Other such mycotoxins, which contain a single nitrogen atom as part of an indole nucleus and also have similarities with much of the indole-diterpene structure,



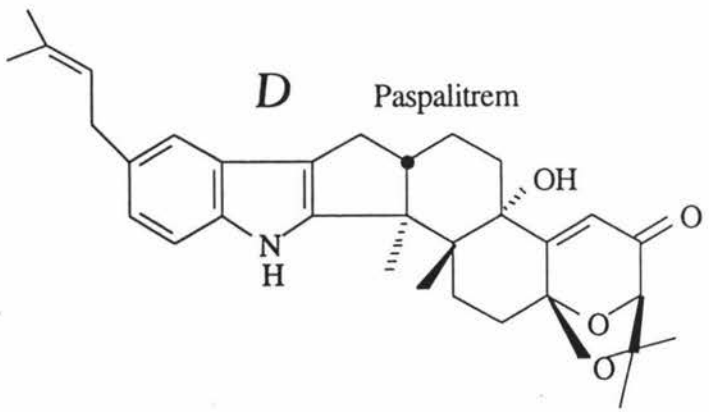

*B* Janthitrem



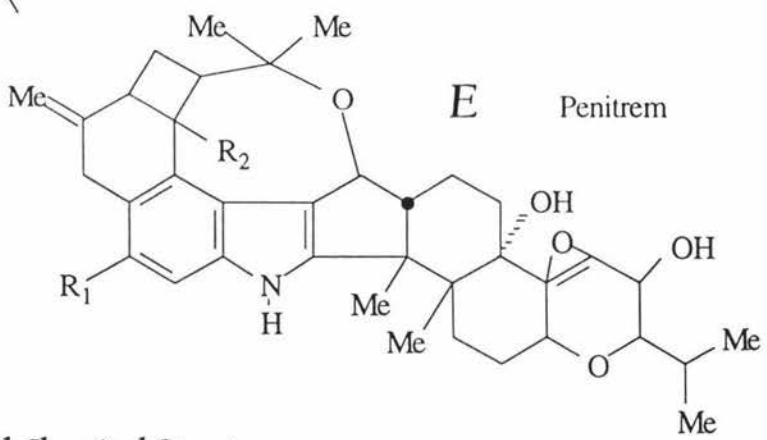

*C* Lolitrem




*A* Paxilline




*D* Paspalitrem




*E* Penitrem

***Fig. 1.1*** *Indole-diterpenoid Chemical Structures*

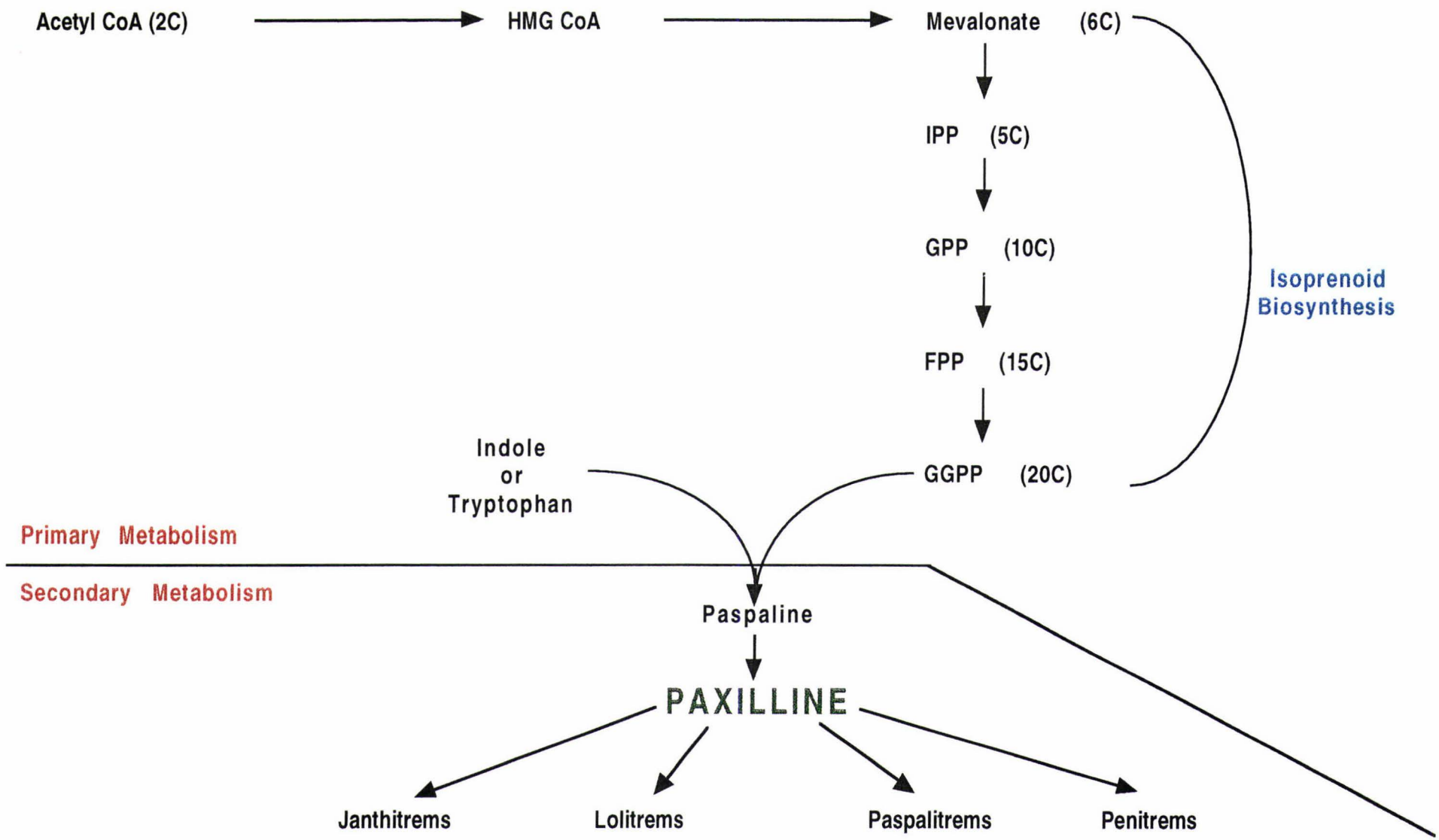



***Fig. 1.2*** *The Proposed Indole-Diterpenoid Pathway*

A schematic diagram showing the boundaries of primary and secondary metabolism within the indole-diterpenoid biosynthetic pathway.

have been grouped into four structural classes (Fig. 1.1 and 1.2): (i) the janthitrems (Gallagher *et al.*, 1980), (ii) the lolitrems (Gallagher *et al.*, 1984), (iii) the paspalitrems (Cole *et al.*, 1977; Dorner *et al.*, 1984) and (iv) the penitrems (de Jesus *et al.*, 1983). Although arising from a diverse range of fungal species, all of these compounds appear to share a common structural backbone of paxilline (Fig 1.1) and are known to be tremorgenic. Paxilline has been readily found in a range of fungi known to produce these indole-diterpenoids (Seya *et al.*, 1986; Weedon and Mantle, 1987), and labelling studies using [$^{14}$C]paxilline have shown that it is incorporated into the penitrems (Mantle and Penn, 1989), supporting the theory that paxilline is a key intermediate for these indole-diterpenoids.

Munday-Finch *et al.* (1996) have proposed a metabolic grid for paxilline production where the initial reaction must be prenylation of indole with geranylgeranyl diphosphate, a 20 membered carbon isoprenoid. Beyond this point, by looking at the individual structural changes between each of the intermediates it maybe possible to predict the type of enzymes that would be associated with this pathway.

## 1.2 *Paxilline Biosynthesis*

### 1.2.1 *Primary Metabolism - The Indole Moiety*

It has long been thought that the indole group of the indole-diterpenoids is incorporated from tryptophan that has lost the aliphatic side chain possibly by tryptophanase activity. This theory has arisen from the fact that tryptophan synthase catalyses the reaction of indole-3-glycerol phosphate to tryptophan via an indole intermediate that is not released (Stryer, 1995). Tryptophan synthase is a bifunctional enzyme with two different domains, an α-subunit (TSA) and a β-subunit (TSB). TSA catalyses the first reaction of indole-3-glycerol phosphate to indole. However, the indole is never released but passes down a 'tunnel' to the β-subunit. It is in TSB that indole is converted to tryptophan and released. The role of tryptophan synthase means that any available indole is sequestered and, therefore, tryptophan could be the only precursor for indole-diterpenoid biosynthesis. However, Frey *et al.* (1997) have recently shown that in addition to the tryptophan synthase α- and β-subunits there is evidence for an independently acting TSA subunit involved in the biosynthesis of two cyclic hydroxamic acids that form part of the defence against insects and microbial pathogens of maize. As labelling studies to date (Acklin *et*

*al.*, 1977; de Jesus *et al.*, 1983; Laws and Mantle, 1989) have not addressed the issue of direct incorporation of indole, it is possible that a TSA-like enzyme is responsible for the indole supply instead of tryptophanase. If such a TSA enzyme was involved with paxilline biosynthesis, it would have to be dedicated to this pathway, therefore, there would be a need for two TSA-like subunits within the genome. One would be associated with primary metabolism and the TSB subunit producing tryptophan while the other, remaining independent and used for secondary metabolism.

### 1.2.2 *Primary Metabolism - The Diterpenoid Moiety*

Isoprenoids are a structurally diverse group of compounds (Chappell, 1995), which can be divided into primary (e.g. sterols, gibberellic acid, carotenoids and those involved in protein prenylation) and secondary metabolites (e.g. monoterpenes, sesquiterpenes and diterpenes). The isoprenoid pathway involves the sequential condensation of three acetyl-CoA units to generate 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA), which is converted to the six carbon mevalonate in an irreversible reaction catalysed by HMG-CoA reductase (Fig. 1.2) Isopentenyl diphosphate (IPP) and its isomer, dimethylallyl diphosphate (DMAPP), are generated by phosphorylation and decarboxylation of mevalonate. It is IPP and DMAPP that are the five-carbon building blocks for all other isoprenoids. Condensation of IPP and DMAPP in a head-to-tail manner generates geranyl diphosphate (GPP); addition of another IPP molecule generates farnesyl diphosphate (FPP) and addition of a third IPP molecule generates geranylgeranyl diphosphate (GGPP). The polymerisation reactions are catalysed by a series of enzymes known as prenyltransferases that are associated with primary metabolism (Fig. 1.2). The group of enzymes responsible for cyclisation of GPP, FPP and GGPP, the terpene, sesqui- and diterpene cyclases, utilise a similar reaction mechanism to the prenyltransferases.

Recent studies in *Neurospora crassa* (Barbato *et al.*, 1996) using RIP (Repeat-Induced Point Mutation) suggests that a complete knockout of the geranylgeranyl diphosphate synthase gene (*al-3*) would be lethal. It appeared that RIP would not introduce more than six point mutations and these mutations were leaky, leaving a partially functional enzyme. Disruption of the geranylgeranyl diphosphate synthase gene (*BST1*) in *Saccharomyces cerevisiae* (Yiang *et al.*, 1995) showed that *BST1* was not essential for the vegetative growth of the yeast cells but, in its absence, growth is impaired at temperatures of 25°C and below. These results suggest that a knockout of geranylgeranyl diphosphate synthase would result not only in the loss of indole-diterpenoid biosynthesis but also would have a large effect on the fungal growth.

### 1.2.3 *Secondary Metabolism - Paxilline Production*

The putative pathway for paxilline biosynthesis (Munday-Finch *et al.*, 1996) has been deduced by the presence of likely chemical intermediates and logical progression of the chemical structures. It is thought that paxilline biosynthesis must begin with the prenylation of indole or tryptophan with GGPP, as well as cyclisation of the 20 carbon isoprenoid unit by a diterpene cyclase or synthase resulting in paspaline. Paspaline is oxidised to paspaline B which undergoes oxidative decarboxylation followed by allylic oxidation to give paxilline. Therefore, based on the need for progressive chemical oxygenation of putative intermediates, one would expect the involvement of oxygenases such as cytochrome P450 monooxygenases. Other expected enzymes include hydrolases and decarboxylases. However, until the pathway has been elucidated biochemically, proposed steps would be only speculative.

Due to the similar structures of paxilline and the other indole-diterpenoid compounds (Fig. 1.1) determination of the paxilline biosynthetic pathway should allow easy isolation of the genes responsible for the other pathways, such as the lolitrems, paspalitrems, penitrems and janthitrems.

## 1.3 *Gene Clusters*

Fungal gene clusters have been broadly defined as the close linkage of two or more genes that participate in a common metabolic or developmental pathway. Clusters commonly found are those of dispensable pathways such as nutrient utilization and the production of secondary metabolites (Keller and Hohn, 1997). These pathways are not required for the normal growth of the organism but are thought to be present for selectional advantages, in particular, as a response to nutrient deprivation or competing organisms. Gene clusters have only recently been found in filamentous fungi. Since the 1980s there has been genetic evidence, using pathway mutants, that established the likely clustering of some genes responsible for aflatoxin biosynthesis (Bennett and Papa, 1988). However, the first molecular evidence of a fungal gene cluster did not arise until 1989, when Hull *et al.* (1989) found the L-proline catabolic pathway of *A. nidulans*.

As many prokaryotic metabolic pathways are clustered and due to similarities seen between some of the fungal and prokaryotic pathways, it is proposed that some fungal-pathway gene clustering may have evolved from horizontal gene transfer from prokaryotes. A good example of this is the isopenicillin N synthase gene from *Streptomyces lipmanii* and *Aspergillus nidulans,* where similarities also include the lack of introns and high GC content (Weigel *et al.*, 1988).

### 1.3.1 *Dispensable Catabolic Pathways*

To date, four dispensable catabolic pathways (quinate, ethanol, proline and nitrate utilisation) are found to be clustered in the filamentous fungus *A. nidulans* (reviewed by Keller and Hohn, 1997). These clusters are in the size range of 12 - 18 kb and appear to contain genes encoding biosynthetic enzymes, transcription factors and transporters. Although some of these pathways are clustered in other fungi, this is not always the case. For example, proline utilisation in *S. cerevisiae* is not clustered, while protein function is, however, still conserved.

### 1.3.2 *Secondary Metabolite - Penicillin*

Since the Second World War there has been a large focus on the production of the β-lactam derivatives, the penicillins and cephalosporins. This would make the penicillin biosynthetic pathway the most well studied of all the secondary metabolites produced by filamentous fungi. Although chemists have tried to establish commercial chemical synthesis of the penicillins, the highly sensitive β-lactam ring does not withstand this, leaving fermentation as the most efficient way of isolating large quantities of penicillin. It was in the 1960s that isolation of linear δ-(α-aminoadipyl)-cysteinyl valine (ACV) from *P. chrysogenum* gave the first evidence that penicillin was a tripeptide made of cysteine, valine and α-aminoadipic acid. In the 1980s the penicillin genes were isolated, with the first evidence of a gene cluster adduced in 1989 by Barredo *et al.*, (1989).

The penicillin gene cluster of *Penicillium chrysogenum* is approximately 20 kb, containing three genes that encode the biosynthetic enzymes, ACV synthetase, isopenicillin N synthetase and acyltransferase. These genes are also known to be clustered in *A. nidulans* and in the cephalosporin-producing *Acremonium chrysogenum.*

### 1.3.3 *Secondary Metabolite - Trichothecenes*

*Fusarium*, a fungus known to cause dry rot in potato tubers, produces a family of toxic sesquiterpenoids, the trichothecenes. These mycotoxins accumulate in agricultural products, are potent inhibitors of eukaryotic protein synthesis and are thought to be involved with plant diseases. Trichothecenes are formed from the primary metabolite, farnesyl diphosphate (FPP) by cyclisation catalysed by trichodiene synthase followed by subsequent modification of trichodiene through a series of oxygenation steps. Isolation of trichodiene synthase showed that it is the first committed step in trichothecene biosynthesis (Beremand, 1987; Hohn and Beremand, 1989). Evidence of a gene cluster was shown by Hohn *et al.* (1993) with the isolation of two overlapping cosmid clones that complemented mutants which blocked the trichothecene pathway. To date, the gene cluster from *Fusarium sporotrichioides* is thought to be ~25 kb with nine genes. Of the six biosynthetic enzymes, two share sequence identity with cytochrome P450 monooxygenases, three share sequence identity with acetyltransferases and the other, based on enzyme activity, is trichodiene synthase. The three remaining genes include a transcription factor that is required for pathway expression (Proctor *et al.*, 1995), an apparent transport protein, and one gene with no known function (Keller and Hohn, 1997).

### 1.3.4 *Secondary Metabolite - Aflatoxin/Sterigmatocystin*

The polyketide mycotoxins, aflatoxin and sterigmatocystin, are potent carcinogens produced by the fungi *Aspergillus flavus, A. parasiticus* (aflatoxin) and *A. nidulans* (sterigmatocystin). These fungi infect corn, cotton and peanuts and are a major threat to the health of humans. Evidence for a gene cluster in *A. parasiticus* was based on the isolation of the closely linked aflatoxin genes, *ver-1* and *nor-1* (Skory *et al.*, 1992), while physical mapping of the *A. flavus apa-2,* an aflatoxin regulatory gene, placed it approximately 8 kb away from *ver-1* (Chang *et al.*, 1993). Keller *et al.* (1994) found evidence that the sterigmatocystin pathway is clustered, when they isolated the *ver*A gene from *A. nidulans*. Two other open reading frames were found alongside the *ver*A gene, one was predicted by sequence identity to encode a P450 monooxygenase (Keller *et al.*, 1994). As with the trichothecene pathway, such genes are expected to be involved with sterigmatocystin and aflatoxin synthesis. The complete sterigmatocystin biosynthetic cluster of *A. nidulans* has been published (Brown *et al.*, 1996) with similarities between the two gene clusters of aflatoxin and sterigmatocystin reviewed by Keller and Hohn

(1997). Although the order of the genes is not identical, the two clusters possess similar gene function and regulatory control. The 60 kb sterigmatocystin cluster contains 25 genes; 16 encoding biosynthetic enzymes, one regulatory gene and eight with unknown functions. Of the 16 biosynthetic enzymes, five show similarity to P450 monooxygenases, four are dehydrogenases and the remaining enzymes include an esterase, an *O*-methyltransferase, a reductase and an oxidase. The boundary of the gene cluster was determined by Northern analysis as fragments that hybridised under both sterigmatocystin-inducing and non-sterigmatocystin-inducing conditions (Brown *et al.*, 1996). At present it is the aim of these groups to knockout each of the genes individually to determine their function.

### *1.3.5 Secondary Metabolite - Melanin*

Melanins are high-molecular-weight, coloured pigments produced by numerous fungi. They give the fungus protection from desiccation, UV irradiation and extreme temperatures as well as cell wall rigidity. They are also thought to be associated with the ability of certain fungal plant pathogens to penetrate host plant cells. Melanins are formed by the oxidative polymerisation of phenolic compounds with the precursors being either polyketide products or tyrosine. A three-gene cluster containing a polyketide synthase, scytalone dehydratase and 1,3,8-trihydroxynaphthalene synthase was identified in the non-pathogenic fungus *Alternaria alternata* (Kimura and Tsuge, 1993). However, with the pathogenic fungi *Colletotrichum lagenarium* and *Magnaporthe grisea,* for example, these genes are not clustered.

### *1.3.6 Secondary Metabolite - Paxilline*

Paxilline, is a known secondary metabolite from *P. paxilli*, that is not required for normal growth and development of this fungus. To date, there is no genetic evidence available, to show that the paxilline biosynthetic genes are clustered. However, as many fungal secondary metabolites are clustered (Section 1.3.1 - 1.3.5) it is also thought the paxilline biosynthetic genes maybe.

## 1.4 *Approaches to Cloning Secondary Metabolite Genes*

### 1.4.1 *Mutagenesis*

UV and chemical mutagenesis have been used for a number of years to isolate a large number of mutants. The advantage of such a system is that a large quantity of random mutants can be quickly isolated with no knowledge of the genes involved. Although easy to perform, there are several drawbacks for using random mutagenesis: (i) there is the possibility of more than one mutation per genome; (ii) the mutations are not tagged in any way making it difficult to isolate the mutated gene; (iii) large numbers of mutants have to be screened to identify the phenotype of interest. However, once mutant has been isolated, several strategies can be used to identify the gene responsible. Beremand (1987) used UV mutagenesis to generate 1000 mutants of *Fusarium sporotrichioides*. Using a microtitre plate to grow the cultures, Beremand (1987) was able to rapidly screen these mutants by a monoclonal antibody to trichothecene. The three mutants identified were characterised chemically for their role in trichothecene biosynthesis. Complementation with a cosmid library of *F. sporotrichioides* genomic DNA containing a hygromycin selectable marker was used to rescue two trichothecene-deficient mutants (Hohn *et al.*, 1993). The cosmids were characterised further to determine the sequences responsible for complementation of the different mutations.

An alternative approach is that of heterologous plasmid tagging by integrative transformation of a plasmid that has no homology with the fungal genome. The plasmid, such as pAN7-1 (Punt *et al.*, 1987), contains a fungal selectable marker (e.g. a gene for hygromycin resistance) and is thought to integrate into the genome of the fungal protoplasts by illegitimate recombination at the site of chromosome nicks or breaks (Fig. 3A). More recently, plasmid tagging has been modified by the addition of a restriction enzyme (Restriction Enzyme Mediated Integration - REMI) during the transformation process (Schiestl and Petes, 1991). A plasmid carrying a fungal selectable marker is linearised with a suitable restriction enzyme and used to transform fungal protoplasts in the presence of the restriction enzyme. Integration of the plasmid occurs at the restriction enzyme target site presumably due to cleavage with the enzyme (Fig. 1.3B). This plasmid tagging method has proven very successful with a number of fungi and is found to increase the transformation frequency and favour the integration of the plasmid as a single copy. REMI was shown to increase the *Cochliobolus heterostrophus* transformation frequency from 6 up to as many as 107 transformants per 30 μg of linear DNA per $10^6$

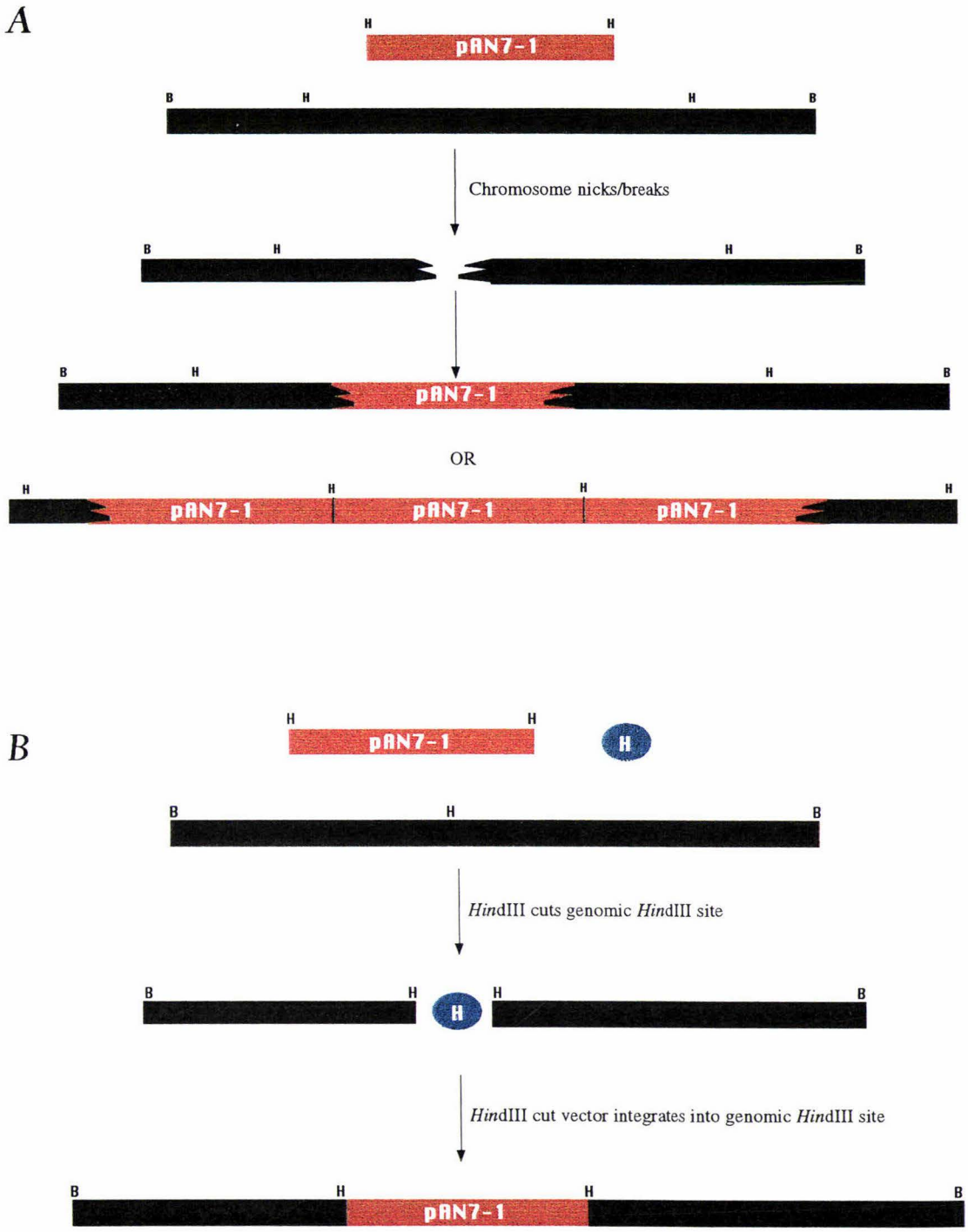



***Fig. 1.3** Plasmid Tagging*

*A* Illegitimate recombination

*B* Restriction Enzyme Mediated Integration (REMI)

The integrating plasmid (pAN7-1) is red. The restriction enzyme is blue. Genomic DNA is black. Abbreviations: B, *Bam*HI. H, *Hin*dIII.

protoplasts when a restriction enzyme was included (Lu *et al.*, 1994). Of the 1310 transformants isolated, two failed to produce detectable T-toxin and virulence to maize was lost. Characterisation of these mutants showed the Tox$^-$ mutation of the two transformants mapped to the expected *Tox1* locus (Lu *et al.*, 1994). DNA recovered from the insertion site of one mutant encodes a 7.6 kb open reading frame that identified a multifunctional polyketide synthase (Yang *et al.*, 1996). Analysis of 1000 REMI insertion mutants of the maize pathogenic fungus *Ustilago maydis* demonstrated that 1-2% of these mutants were unable to induce symptoms when tested *in planta*. Characterisation of two of these mutants showed that the phenotype is linked to the insertion event (Bölker *et al.*, 1995).

With all mutagenesis techniques there is a requirement to screen large numbers of mutants to find the phenotype of interest. Therefore, an observable phenotype is very useful, for example, spore pigmentation (Itoh and Scott, 1994) or an ability to produce carotenoids such as those that give *N. crassa* its orange colour (Barbato *et al.*, 1996). However, for most mutants an observable phenotype is not available. More time-consuming screens involve *in planta* assays for pathogenicity (Bölker *et al.*, 1995) or extraction of the compound of interest for analysis by TLC (Hodges *et al.*, 1994). Monoclonal-based ELISA are useful for quick screening of culture filtrates for compounds of interest (Beremand, 1987). Such techniques have a quick assay system. However, the antibody needs to be reasonably specific for the compound being observed.

The advantage of plasmid tagging over the other mutagenesis techniques is the ability to rescue the tagged gene via plasmid rescue. Of course, there is still the labour intensive screen for transformants with the appropriate phenotype, as is also required for UV mutagenesis.

### 1.4.2 *Enzyme Isolation*

When an enzyme has a known biochemical activity that can be assayed it is possible to isolate the protein. A number of secondary metabolite genes have been found as a result of protein purification. Tricodiene synthase, the enzyme that catalyses the isomerisation and cyclisation of FPP to tricodiene, was isolated from *Fusarium sporotrichioides*. Antibodies to this enzyme were made and used to screen a λgt11 expression library to *F. sporotrichioides*. The resulting cloned DNA was sequenced and compared to the known protein sequence from CNBr cleaved fragments, confirming that the gene was tricodiene synthase. An internal fragment from the cDNA was used to isolate a cosmid containing the complete gene sequence. Expression of this gene and biochemical analysis confirmed the

tricodiene synthase activity (Hohn and Beremand, 1989). Tsai *et al.* (1995) used a slightly different approach. The isolated enzyme, dimethylallyl tryptophan synthase (DMAT), was first purified from mycelia of *Claviceps purpurea* (Gebler and Poulter, 1992). This enzyme is responsible for the alkylation of L-tryptophan by dimethylallyl diphosphate and is the first pathway-specific step of ergot alkaloid biosynthesis. Partial sequence of CNBr fragments was used to design degenerate primers that hybridised to *C. purpurea* genomic DNA. PCR was used with these primers and the nucleotide sequence compared to the available DMAT peptide sequence. The complete DMAT gene was cloned and sequenced, and its function confirmed by expression in a yeast background.

For the above method to be useful, protein sequence data are required. These are not always available and enzyme purification can be a lengthy process even with a biochemical assay. With the sequence databases increasing rapidly in size everyday, it is often possible to obtain and compare known protein sequences. If there is sufficient identity between the proteins, degenerate primers can be designed and used for PCR. The drawback of this system is finding regions of the sequence that contain enough identity with amino acids that have low codon usage. Unfortunately many proteins have similar activity but sequence conservation can be limited to a catalytic domain or an active site.

### 1.4.3 *Heterologous probing*

*Aspergillus* species are known to produce the similar compounds aflatoxin and sterigmatocystin. As a number of genes have been isolated from the aflatoxin-producing fungi, *A. parasiticus* and *A. flavus*, it is possible to use these cloned fragments for the isolation of homologues in *A. nidulans*. The *ver-1* gene from *A. parasiticus* is responsible for the enzymatic conversion of versicolorin A to sterigmatocystin (Skory *et al.*, 1992). Heterologous hybridisation was used to isolate *ver*-1 homologue from *A. nidulans* (Keller *et al.*, 1994). The hybridising fragment was subcloned from a cosmid and sequenced.

This technique has also been used for the isolation of dothistromin genes. Dothistromin is a polyketide toxin produced in pine by the fungus *Dothistroma pini.* As the chemical structure of dothistromin is similar to aflatoxin, similar enzymes maybe responsible for its production. Heterologous probing with the *ver-1* gene from *A. parasiticus* was used to isolate the *D. pini* homologue from a genomic library (Gillman and Bradshaw, personal comm.). Sequence data show that the *D. pini ver* gene has 75% identity to the *A. parasiticus* gene (Monahan and Bradshaw, personal comm.).

### 1.4.4 *Differential screening*

In recent years, there has been a number of RNA-based techniques established for looking at differentially expressed genes. Different culture conditions (e.g. temperature, time and nutrient sources) can induce or repress gene expression, these conditions are then used for RNA isolation. The earliest of these techniques was subtractive hybridisation. Feng *et al.* (1992) used growth phase, carbon source and temperature to distinguish differential expression of genes related to aflatoxin biosynthesis. The three labeled cDNA probes were hybridised successively to the same membranes hosting the DNA library. When compared, 0.6% of the 20 000 colonies were found to differentially hybridise to the three probes. These clones were screened against Northern blots to further prove their expression pattern.

Another differential screening method is that of differential display (DD) (Liang and Pardee, 1992). This is a PCR-based technique that compares differences in mRNA content between two different states. Total RNA isolated from cultures grown under different physiological conditions is amplified by reverse transcriptase with poly-T tailed anchored primers. These primers anneal to the poly-A tail of the mRNA and are anchored at the 3' end by A, C or G nucleotides. The second round of PCR uses the anchored poly-T primer with a primer of arbitrary sequence. This reduces the pool of represented mRNA so the fragments can be separated on a denaturing gel. The bands from the different growth conditions are compared to find fragments that are differentially expressed. These bands are isolated, reamplified and used to screen Northern blots to confirm their expression pattern.

### 1.4.5 *Proteonome Analysis*

There have been great advancements with two-dimensional (2D) gel electrophoresis of proteins, and with NMR and mass spectrophotometry for obtaining a protein sequence from a small sample. This has resulted in a revival in proteonome analysis, where the proteins of two physiological states are compared by 2D gel electrophoresis. Any proteins that are differentially expressed can be sequenced by mass spectrophotometry directly from the gel. It is also possible to determine protein function by comparison of 2D gel libraries of known proteins.

## 1.5 *Genetic Criteria to Confirm Functionality of the Isolated Gene.*

Once a potential gene is isolated, there needs to be a way of confirming its actual physiological role. In sexual fungi, crosses can be used to determine whether the phenotype and marker co-segregate. However, this approach cannot be used with asexual fungi and, therefore, other methods are required. In many cases, complementation can be used when a gene has been knocked out. When replacing it with a functional gene, the wild-type phenotype should be restored. This is not always possible due to the lack of available selectable markers for filamentous fungi. When complementation is not possible, homologous recombination can be used to knockout the gene of interest in a wild-type strain, resulting in a recessive phenotype. This approach can be lengthy as the double crossover integration event required for a true knockout occurs at a low frequency and is locus dependent (Bird *et al.*, 1997).

## 1.6 *Cloning of the Paxilline Biosynthetic Genes in Penicillium paxilli by Plasmid Integration*

As the genes of many fungal secondary metabolite pathways are arranged in clusters, it would be of no surprise to find this the case for the paxilline biosynthetic genes. To isolate such genes from *P. paxilli* the approach initiated was that of heterologous plasmid integration using a monoclonal-based competitive ELISA (Garthwaite *et al.*, 1993) to screen for the loss of paxilline production. If the genes responsible for paxilline biosynthesis are found in a cluster, random plasmid integration into one of these genes, would knock out paxilline production. Therefore, isolation of sequences flanking the plasmid integration site and the surrounding locus would result in the cloning of the clustered genes.

Itoh *et al.* (1994) used pAN7-1 (Appendix A1.1) to isolate 600 tagged transformants of *P. paxilli*. A sub-group of these transformants was checked for stability, frequency and randomness of integration. Up to 78% of the transformants contained the plasmid at a single site, and 50% of these transformants were present as a tandem repeat. The integration of the plasmid into the genome did not appear to be site specific. In contrast, work with *Aspergillus nidulans* using heterologous plasmid integration showed site-selectivity (Diallinas and Scazzocchio, 1989; Tilburn *et al.*, 1990). Within the 600

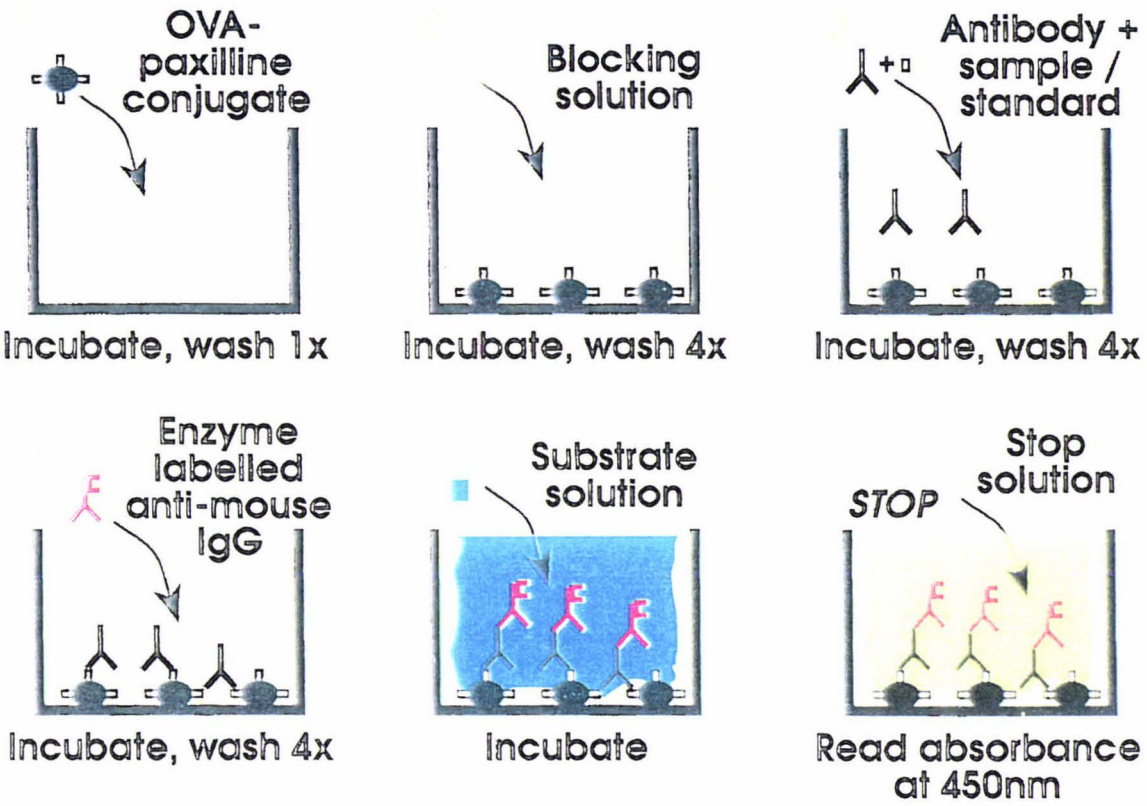



***Fig. 1.4*** *The Competitive ELISA for Paxilline*

Courtesy of Ian Garthwaite, Toxinology and Food Safety Research Group, Ruakura, Hamilton.

pAN7-1 tagged transformants, 0.83% had a characteristic brown spore phenotype. Isolation and characterisation of mutants in this spore pigmentation locus showed that the integration of pAN7-1 had caused extensive deletions through a common region of the genome (Itoh and Scott, 1994). All of the above mutants were also screened by the cELISA (Fig. 1.4) for their ability to produce paxilline in culture. The antibody used in this assay recognises both paxilline and an array of compounds structurally related to paxilline. One mutant, YI-20, had a paxilline negative phenotype as determined by cELISA and HPLC. Characterisation of this mutant by Southern analysis (Fig 1.5) showed that the integrating plasmid, pAN7-1, was present at a single site but in a head-to-tail tandem repeat. Hybridisation of genomic digests of YI-20 DNA with pAN7-1, shows bands present at 6.8 kb in the *Bam*HI (lane 2) and *Hin*dIII (lane 3) digests, indicative of a tandem repeat (Itoh, unpublished data).

## 1.7 *Aim*

The aim of this thesis is to characterise the Pax$^-$ mutant, YI-20, with respect to its integration of pAN7-1 and to establish whether plasmid integration was responsible for the phenotype observed. The genomic sequence surrounding the pAN7-1 integration will be isolated. These sequences will be used to make a replacement construct with the intent to generate additional paxilline-negative mutants via homologous recombination. The resulting mutants will be characterised for their ability to produce paxilline and for their integration events.

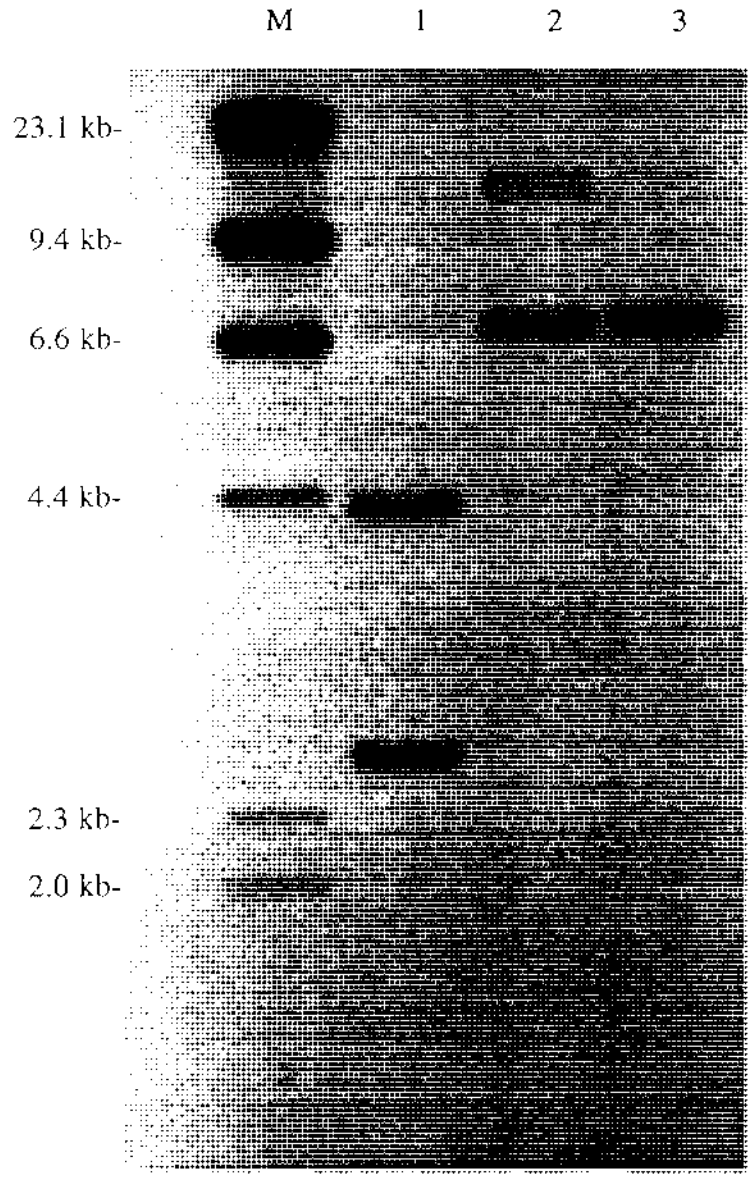



***Fig. 1.5*** *Southern analysis of the integration of pAN7-1 in YI-20*

Southern analysis of YI-20 genomic DNA digested with *Eco*RI (lane 1), *Bam*HI (lane 2), *Hin*dIII (lane 3), hybridised with pAN7-1. Lane M, λ *Hin*dIII.

# *Chapter Two*
# *Materials and Methods*

## 2.1 *Fungal and Bacterial Strains, λ Clones and Plasmids*

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

## 2.2 *Growth of Cultures*

### 2.2.1 *Fungal Spore Suspensions*

*P. paxilli* cultures were grown at 22°C on potato dextrose (PD agar, Section 2.3.5) agar plates for 5 - 7 days. Squares of agar were cut from these plates and resuspended in 0.01% (v/v) triton X-100 (Section 2.4.1). The spore concentration was determined using a haemocytometer slide. Spore suspensions were maintained either at 4°C or as suspensions in 10% (v/v) glycerol (Section 2.4.1) at -70°C.

### 2.2.2 *Fungal Spore Purification*

Single spore isolates of *P. paxilli* were obtained by streaking a loop full of spores suspended in 0.01% (v/v) triton X-100 (Section 2.4.1) onto PD plates (Section 2.3.5) and growing for 5 - 7 days. A single colony was restreaked and spore suspensions were then made from the replating.

### 2.2.3 *Fungal Growth Conditions*

When requiring mycelia for DNA analysis and protoplasting, 25 ml of PD broth (Section 2.3.5) was inoculated with 5 x $10^6$ spores and grown for 40 h at 22°C, shaking at 200 rpm.

Cultures to be assayed for paxilline were grown for 5 - 7 days at 28°C in 25 ml of CDYE broth (Section 2.3.1) inoculated with 1 x $10^7$ spores, shaking at 250 rpm.

### *2.2.4 Bacterial Growth Conditions*

*E. coli* cultures were grown at 37°C in LB broth or on LB agar plates (Section 2.3.4). Cultures were stored at -70°C in 50% (v/v) glycerol.

***Table 2.1*** *Fungal and Bacterial Strains, λ clones and plasmids.*

| Strain | Relevant characteristics | Source or reference |
|---|---|---|
| Fungal strains | | |
| *P. paxilli* | | |
| PN2013 | Wild-type Pax$^+$ | Itoh *et al.*, 1994 |
| CY-2 | Pax$^-$Hyg$^R$ | This study |
| CY-35 | Pax$^+$Hyg$^R$ | This study |
| CY-102 | Pax$^-$Hyg$^R$ | This study |
| YI-20 | Pax$^-$Hyg$^R$ | This study |
| YI-186 | Pax$^+$Hyg$^R$ | Itoh and Scott, 1997 |
| Bacterial strains | | |
| *E coli* | | |
| KW251 | *F$^-$supE44 galK galT22 metB1 hsdR2 mcrB1 mcrA* [*argA81* :Tn10]*recD1014* | Promega Corp. |
| XL-1 | *supE44 hsdR17 recA1 endA1 gyrA46 thi relA1 lac$^-$* F'[*proAB$^+$ lacI$^q$ lacZ*ΔM15 Tn10 (Tet$^R$)] | Bullock *et al.*, 1987 |
| PN1199 | HB101/pAN7-1 | Murray *et al.*, 1992 |
| PN1375 | XL1/pPN1375 | This study |
| PN1385 | XL1/pPN1385 | This study |
| PN1387 | XL1/pPN1387 | This study |
| PN1389 | XL1/pCWHyg1 | This study |
| PN1406 | XL1/pPN1406 | This study |
| PN1408 | XL1/pPN1408 | This study |
| PN1414 | XL1/pPN1414 | This study |
| PN1418 | XL1/pPN1418 | This study |
| PN1421 | XL1/pPN1421 | This study |
| PN1422 | XL1/pPN1422 | This study |
| Lambda Clones | | |
| λCY1 | λGEM11 clone from *P. paxilli* | This study |
| λCY3 | λGEM11 clone from *P. paxilli* | This study |

***Table 2.1*** *continued*

| Plasmid | Relevant characteristics | Source or reference |
|---|---|---|
| Plasmids | | |
| pAN7-1 | 6.8 kb $Hyg^R Amp^R$ | Punt *et al.*, 1987 |
| pCWHyg1 | 4.8 kb $Hyg^R Amp^R$ | Cathy Wasmann University of Arizona |
| pPN1375 | 6.9 kb plasmid rescued from *P. paxilli* YI-20 $Hyg^R Amp^R$ | This study |
| pPN1385 | 1.7 kb *Eco*RI fragment from λCY1 in pUC118 $Amp^R$ | This study |
| pPN1387 | 2.8 kb *Hin*dIII fragment from λCY1 in pUC118 $Amp^R$ | This study |
| pPN1406 | 2.5 kb *Sst*I fragment from λCY1 in pUC118 $Amp^R$ | This study |
| pPN1408 | 4.5 kb *Sst*I fragment from λCY1 in pUC118 $Amp^R$ | This study |
| pPN1414 | pUC118 with 2.8 kb *Hin*dIII fragment from pPN1387 and 2.3 kb *Hin*dIII/*Sal*I $Hyg^R$ from pCWHyg1 | This study |
| pPN1418 | pPN1414 containing 2.5 kb *Sst*I fragment from pPN1406 | This study |
| pPN1421 | 3.3 kb *Sst*I fragment from λCY1 in pUC118 $Amp^R$ | This study |
| pPN1422 | 3.5 kb *Sst*I fragment from λCY1 in pUC118 $Amp^R$ | This study |
| pUC118 | 3.2 kb $Amp^R$ | Vieira and Messing, (1987) |

## 2.3 *Media*

All media were prepared with MilliQ water and sterilised at 121°C for 15 min, unless stated otherwise. Where appropriate, media were supplemented as in Section 2.3.8.

### 2.3.1 *CD + Yeast Extract Media (CDYE)*

CDYE media contained 3.34% (w/v) Czapek Dox (CD, Oxoid) and 0.5% (w/v) yeast extract.

### 2.3.2 *Complete Media (CM)*

Complete media (Oliver *et al.*, 1987) contained (g/l); Czapek Dox (Oxoid), 33.4; yeast extract 1.0; potato dextrose broth 12.0; mycological peptone, 1.0; casein hydrolysate, 1.0; 0.8 M sucrose and agar, 15.0. The pH was adjusted to 6.5 with NaOH and then autoclaved at 110°C.

### 2.3.3 *Complete Media Top Agar*

Complete media top agar (Oliver *et al.*, 1987) was prepared as above (Section 2.3.2) but contained 0.8% agar.

### 2.3.4 *LB Media*

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

### 2.3.5 *Potato Dextrose Media (PD)*

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

### 2.3.6 *SOC Media*

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

### 2.3.7 *Top Agarose*

Top agarose contained 1% (w/v) tryptone (Difco), 0.5% (w/v) NaCl, 0.8% (w/v) agarose 15 (BDH). Before use, the top agarose was supplemented with $MgSO_4.7H_2O$ (Section 2.4.1) to a final concentration of 10 mM.

### 2.3.8 *Media Supplements*

| Supplement | Stock concentration | Final concentration |
|---|---|---|
| Ampicillin | 100 mg/ml | 100 µg/ml |
| Hygromycin[a] | 50 mg/ml | 200, 100, 20 µg/ml |
| IPTG | 24 mg/ml | 38 µg/ml |
| X-gal[b] | 20 mg/ml | 32 µg/ml |

[a]Boehringer Mannheim.

[b]Dissolved in dimethyl formamide.

## 2.4 *Common Buffers and Solutions*

### 2.4.1 *Stock Solutions*

| Stock | Concentration | pH |
|---|---|---|
| Ethanol | 70% (v/v) | |
| Ethidium bromide | 10 mg/ml | |
| Glycerol | 10%, 50% (v/v) | |
| Lithium chloride | 4 M | |
| Maltose | 20% (w/v) | |
| $MgSO_4.7H_2O$ | 1 M | |
| NaCl | 1 M | |
| SDS | 10% (w/v) | |
| Triton X-100 | 0.01% (v/v) | |
| $Na_2EDTA$ | 250 mM | 8.0 |
| Sodium acetate | 3 M | 7.0, 4.5 |
| Tris | 1 M | 8.0, 7.6 |

## 2.4.2 *Common Solutions*

### *10 x Denhardts*

10 x Denhardts (Southern, 1975) contained (per litre): 50 ml of 1 M HEPES (Sigma), pH 7.0; 150 ml 20 x SSC (Section 2.4.2); 18 mg phenol-extracted herring sperm DNA (Sigma); 1.0 g SDS; 20 mg *Escherichia coli* tRNA; 2 g ficoll (Sigma 70); 2 g bovine serum albumin (Sigma); 2 g polyvinylpyrrolidone (Sigma PVP-10).

### *Acrylamide Mix*

Acrylamide mix for sequencing gels was prepared as follows: 288 g urea; 34.2 g acrylamide; 1.8 g bis-acrylamide; made up to 500 ml and deionised with 5 g amberlite MB-3 (Sigma) for at least 30 min; filtered through a porous sintered glass funnel and made up to a final volume of 600 ml with 10 x sequencing TBE (60 ml, Section 2.4.2) and MilliQ water.

### *Alkaline Lysis Solutions*

| | |
|---|---|
| *Solution I* | 50 mM glucose, 25 mM Tris-HCl, 10 mM $Na_2EDTA$, pH 8.0. |
| *Solution II* | 0.2 M NaOH, 1% (w/v) SDS. |
| *Solution III* | 29.44 g potassium acetate, 11.5 ml glacial acetic acid/100 ml. |

### *Blotting Solutions*

| | |
|---|---|
| *Solution 1* | 0.25 M HCl. |
| *Solution 2* | 0.5 M NaOH, 0.5 M NaCl. |
| *Solution 3* | 2.0 M NaCl, 0.5 M Tris-HCl, pH 7.4. |
| *20 x SSC* | 5 M NaCl, 0.3 M sodium citrate. |
| *5 x SSC* | 1.25 M NaCl, 0.075 M sodium citrate. |
| *2 x SSC* | 0.5 M NaCl, 0.03 M sodium citrate. |

### *Dialysis Tubing*

Dialysis tubing was prepared (Sambrook *et al.*, 1989) by boiling in 2% (w/v) sodium carbonate, 1 mM $Na_2EDTA$ (pH 8.0), rinsed well in MilliQ water, then boiled for a further 10 min in 1 mM $Na_2EDTA$ (pH 8.0). The dialysis tubing was stored at 4°C.

*DIG solutions*

| | |
|---|---|
| *DIG Hybridisation Buffer* | 5 x SSC (Section 2.4.2), 0.1% (w/v) sodium lauroylsarcosine, 0.02% (w/v) SDS and 1.0% (w/v) blocking reagent (Boehringer). |
| *DIG Buffer One* | 150 mM NaCl, 100 mM maleic acid (pH 7.5). |
| *DIG Buffer Two* | 150 mM NaCl, 100 mM maleic acid (pH 7.5), 1.0% (w/v) blocking reagent (Boehringer). |
| *DIG Buffer Three* | 100 mM Tris-HCl (pH 9.5), 100 mM NaCl. |

*DNase (RNase Free)*

Deoxyribonuclease I (10 mg/ml) was made up in 0.15 M NaCl and 50% (v/v) glycerol, and stored at -20°C.

*Lysozyme*

Lysozyme (10 mg/ml) was made up in 10 mM Tris-HCl (pH 8.0) and stored at -20°C in small aliquots.

*Protoplasting Solutions*

All solutions required for protoplasting were sterile.

| | |
|---|---|
| *OM Buffer* | 1.2 M $MgSO_4.7H_2O$, 10 mM $Na_2HPO_4$, the pH was adjusted to 5.8 with 100 mM $NaH_2PO_4.2H_2O$. |
| *ST Buffer* | 0.6 M sorbitol, 100 mM Tris-HCl (pH 8.0). |
| *STC Buffer* | 1 M sorbitol, 50 mM $CaCl_2$, 50 mM Tris-HCl (pH 8.0). |
| *PEG Buffer* | 40% (w/v) PEG 4000, 50 mM $CaCl_2$, 1 M sorbitol, 50 mM Tris-HCl (pH 8.0). |
| *GMB Buffer* | 0.9 M sorbitol, 0.125 M $Na_2EDTA$ (pH 7.5). |
| *LMP in GMB* | 1.4% (w/v) low melting point agarose (Sigma) in GMB buffer. |
| *SE Buffer* | 2% (w/v) SDS, 0.25 M $Na_2EDTA$ (pH 8.0). |
| *10 x ET* | 10 mM Tris, 0.5 M $Na_2EDTA$ (pH 8.0). |
| *1 x ET* | 1 mM Tris, 0.05 M $Na_2EDTA$ (pH 8.0). |

*RNase (DNase Free)*

RNaseA (Sigma) was prepared at 10 mg/ml in 10 mM Tris-HCl (pH 7.5), 15 mM NaCl, heated to 100°C for 15 min and allowed to cool slowly to room temperature. RNaseA was stored in small aliquots at -20°C. When used in an enzyme digest, the RNase was further diluted to 0.2 mg/ml.

*SDS Loading Dye*

SDS loading dye contained 1% (w/v) SDS, 0.02% (w/v) bromophenol blue, 20% (w/v) sucrose and 5 mM $Na_2EDTA$ (pH 8.0).

*SM Buffer*

SM buffer contained 100 mM NaCl, 8 mM $MgSO_4.7H_2O$, 50 mM Tris-HCl (pH 7.5) and 0.01% (w/v) gelatin.

*SSC-Saturated Isopropanol*

Equal volumes of 20 x SSC (Section 2.4.2) and isopropanol were stirred overnight. The top layer was used for extracting ethidium bromide from DNA solutions.

*TAE Buffer*

At 1 x concentration TAE buffer contained 40 mM Tris-acetate, and 2 mM $Na_2EDTA$, pH 8.5.

*TBE Buffer*

At 1 x concentration TBE buffer contained 89 mM Tris, 89 mM boric acid, and 2.5 mM $Na_2EDTA$, pH 8.2.

*TBE Sequencing Buffer*

At 1 x concentration TBE buffer contained 134 mM Tris, 45 mM boric acid, and 2.5 mM $Na_2EDTA$, pH 8.8.

*Tris/EDTA Buffers (TE and TES)*

| | |
|---|---|
| *TE 10/1* | 10 mM Tris-HCl, 1 mM $Na_2EDTA$, pH 8.0. |
| *TE 10/0.1* | 10 mM Tris-HCl, 0.1 mM $Na_2EDTA$, pH 8.0. |
| *TES 10/1/100* | 10 mM Tris-HCl, 1 mM $Na_2EDTA$, 100 mM NaCl, pH 8.0. |

*Tris-equilibrated Phenol*

Tris-equilibrated phenol was supplied by Amersham. 8-Hydroxyquinoline was added to give a final concentration of 0.1% (w/v). The equilibrated phenol was stored at 4°C.

## 2.5 *DNA Isolation*

### 2.5.1 *"Maxiprep" Isolation of DNA from Fungal Cultures*

DNA was extracted from fungal cultures using the method of Byrd *et al.* (1990). Freeze-dried mycelia (100 mg, Section 2.2.3) were ground to a fine powder under liquid nitrogen and suspended thoroughly in 10 ml of extraction buffer (150 mM $Na_2EDTA$, 50 mM Tris-HCl (pH 8.0), 1% (w/v) sodium lauroylsarcosine and 2 mg/ml proteinase K, Boehringer). The solution was centrifuged at 2 000 *g* for 10 min (4°C). The supernatant was incubated at 37°C for 20 min, extracted with an equal volume of phenol and spun at 20 200 *g* for 15 min (4°C). The aqueous phase was extracted in a similar manner with equal volumes of phenol/chloroform and finally chloroform. The aqueous phase was centrifuged at 25 000 *g* for 20 min and the DNA was precipitated with one volume of isopropanol at 4°C for 10 min. After centrifugation (8 000 *g*, 10 min, 4°C) the pellet was washed with 70% ethanol (Section 2.4.1), dried and resuspended in 500 µl of TE 10/0.1 (section 2.4.2).

### 2.5.2 *"Miniprep" Isolation of DNA from Fungal Cultures*

DNA was isolated on a small scale using modifications of the Yoder (1988) method as described. Freeze-dried mycelia (10 mg, Section 2.2.3) were ground up with liquid nitrogen in an Eppendorf tube and homogenised with 500 µl of buffer (100 mM LiCl, 10 mM $Na_2EDTA$, 10 mM Tris-HCl (pH 7.4), 0.5% (w/v) SDS). This mixture was extracted with Tris-equilibrated phenol (500µl, Section 2.4.2) and spun for 15 min. The aqueous phase was extracted with chloroform, and the DNA precipitated with 800 µl of

95% ethanol. The precipitated DNA was washed with 70% ethanol, dried and resuspended in 50 µl MilliQ $H_2O$.

### 2.5.3 *Extraction of DNA from Bacterial Cultures by the Rapid Boil Method*

Plasmid DNA from *E. coli* was isolated by the method of Holmes and Quigley (1981). An overnight culture (1.5 ml, Section 2.2.4) was spun down, and the cells resuspended in 350 µl of STET buffer (8% (w/v) sucrose, 5% (v/v) Triton X-100, 50 mM Tris-HCl (pH 8.0), 50 mM $Na_2EDTA$). Lysozyme was added (25 µl, section 2.4.2), the solution boiled for 40 sec and spun for 10 min. After removal of the gelatinous pellet the DNA was precipitated with an equal volume of isopropanol at -20°C for 20 min. The DNA was spun for 10 min, washed with 70% ethanol, dried and resuspended in 50 µl of MilliQ $H_2O$.

### 2.5.4 *Extraction of DNA from Bacterial Cultures by the Alkaline Lysis Method*

Plasmid DNA that would require sequencing was prepared by the method of Sambrook *et al.* (1989). An overnight culture (1.5 ml, Section 2.2.4) was spun down, the cells resuspended in 100 µl of solution I (Section 2.4.2), followed by 200 µl of solution II (Section 2.4.2); the contents of the tube were mixed thoroughly by inversion, 150 µl of solution III (Section 2.4.2) was added and the mixture vortexed in an inverted position. The solution was incubated on ice for 3 - 5 min and spun for 5 min. The supernatant was extracted with an equal volume of Tris-equilibrated phenol (Section 2.4.2) and spun for 5 min. The aqueous phase was extracted with equal volumes of phenol/chloroform, then with an equal volume of chloroform. The DNA was precipitated with 2 volumes of 95% ethanol, spun for 5 min, washed with 70% ethanol, dried and resuspended in 50 µl of MilliQ $H_2O$.

### 2.5.5 *Extraction of DNA from Bacterial Cultures by the Caesium Chloride Method*

DNA from *E. coli* cells harbouring the plasmid of interest was isolated by an alkaline lysis/CsCl method modified from Ish-Horowitz and Burke (1981). A 5 ml overnight culture (Section 2.2.4) was used to inoculate 500 ml of LB broth (section 2.3.4) and grown overnight at 37°C shaking at 200 rpm. The cells were harvested by centrifugation at 10 000 *g* for 10 min (4°C), washed with TE (10/1, section 2.4.2) and re-harvested as

above. The cells were resuspended in 30 ml of solution I (Section 2.4.2) containing 150 mg lysozyme (Boehringer) and incubated at room temperature for 10 min. A 60 ml aliquot of solution II (Section 2.4.2) was added, mixed by inversion and the mixture incubated on ice for 10 min. A 45 ml aliquot of solution III (Section 2.4.2) was then added, mixed by inversion and the mixture incubated for a further 10 min on ice, and centrifuged for 10 min at 10 000 *g* (4°C). Isopropanol (0.6 vol) was added to the supernatant and the mixture incubated on ice for 20 min. The DNA was pelleted by centrifugation at 23 500 *g* for 10 min (4°C), washed with 70% ethanol, dried and resuspended in 7.5 ml of TE (10/1, section 2.4). Caesium chloride (CsCl) was added to the DNA solution in a ratio of 1.05 g CsCl/ml of DNA solution, followed by ethidium bromide (section 2.4.1) in a ratio of 75 µl ethidium bromide/ml of DNA/CsCl solution. The solution was mixed well, left overnight at 4°C then spun at 17 000 *g* for 10 min (4°C). The supernatant was checked for a refractive index of 1.3920 and adjusted with either TE (10/1, section 2.4.2) or CsCl if necessary. The DNA was spun at 166 000 *g* for 5 h (18°C) and the resulting band corresponding to plasmid DNA removed by an 18-gauge hypodermic needle and syringe. The ethidium bromide was removed by mixing with an equal volume of SSC-saturated isopropanol (section 2.4.2), repeating the extraction until the DNA solution became clear. The CsCl was removed by dialysis (dialysis tubing, Section 2.4.2) against TES (10/1/100, section 2.4.2). The resulting DNA was quantified by spectrophotometric analysis (Section 2.6.3) and stored at 4°C.

### 2.5.6 *Purification of DNA from Phage λ*

Phage plated (Section 2.12) for confluent lysis on LB (Section 2.3.4) agarose 15 (BDH) were overlayed with 5 ml of SM buffer (Section 2.4.2) and left overnight at 4°C. The resulting lysate was collected and the bacterial debris removed by centrifugation at 3000 *g* for 10 min (4°C). DNase and RNase (Section 2.4.2) were added to the lysate to give a final concentration of 1 µg/ml and incubated at 37°C for 30 min. A 5 ml aliquot of PEG 6000 solution (20% (w/v) PEG 6000 in 2 M NaCl) was added, the solution mixed and incubated on ice for 1 h. The phage was pelleted by centrifugation at 5800 *g* for 30 min (4°C); the pellet was resuspended in 0.5 ml of SM buffer with 5 µl of 10% (w/v) SDS (Section 2.4.1) and 10 µl of 250 mM $Na_2EDTA$ (Section 2.4.1), and incubated at 68°C for 15 min. The phage coat and proteins were removed by extracting with an equal volume of phenol (Section 2.4.2). The aqueous phase was further extracted with equal volumes of phenol/chloroform and finally with an equal volume of chloroform. The DNA was precipitated by the addition of two volumes of 95% ethanol and 50 µl of 3 M sodium acetate (Section 2.4.1), incubated on ice for 30 min and spun for 10 min in an Eppendorf

centrifuge. The DNA pellet was washed with 70% ethanol and resuspended in 50 µl of TE (10/0.1, Section 2.4.2).

## 2.6 *DNA Quantification*

DNA was quantified by one of three methods as required.

### 2.6.1 *Minigel Method*

A 2µl aliquot of sample DNA was compared to a series of standards of known DNA concentration by separation on a minigel (Section 2.10). Three different types of concentration standards are available and described below.

| Type of standard | Range of standard | Used for |
|---|---|---|
| Uncut phage λ | 10 ng - 200 ng | Concentration of genomic DNA |
| Linearised pUC118 | 2.5 ng - 20 ng | Concentration of fragments in the 1.5 kb - 10 kb size range |
| Low Mass Ladder (BRL) | 10 ng - 100 ng | Concentration of small fragments less than 4 kb size. |

### 2.6.2 *Fluorometric Quantitation*

For impure DNA samples, or pure samples of low concentration, DNA was quantitated on a Hoefer Scientific TKO 100 fluorometer. DNA was quantitated in a dye solution containing 1 x TNE (10 mM Tris-HCl (pH 7.4), 1 mM $Na_2EDTA$, and 100 mM NaCl) and 0.1 µg/ml of Hoechst 33258 dye. The scale of the fluorometer was zeroed with dye solution until a steady state was obtained, then set to 100 using 2 µl of a 100 µg/ml calf thymus DNA added to 2 ml of dye solution. Once the scale was reliably set, 2 µl of sample DNA was added to 2 ml of dye solution, giving the concentration of the sample DNA solution in ng/µl.

### 2.6.3 *Spectrophotometric Quantitation*

Spectrophotometric DNA quantitation was used for DNA isolated by CsCl extraction (Section 2.5.5), as this was the only DNA pure enough for this method. With the sample in quartz cuvettes the concentration of DNA was determined from absorbance readings at 260 nm and 280 nm. The reading at 260 nm allowed calculation of the concentration of nucleic acid present in the sample; an absorbance of 1 corresponds to approximately 50 μg/ml of double-stranded DNA. The absorbance ratio, $A_{260}/A_{280}$, was used as an estimate of DNA purity, as pure DNA has a ratio of 1.8.

## 2.7 *Restriction Endonuclease Digestion of DNA*

Restriction endonuclease (RE) digests were carried out in the commercial buffer recommended by the manufacturer. RE digests contained an excess of RE enzyme (within the limits of star activity) and were performed in a water bath at temperatures recommended by the manufacturer. A small aliquot of digested DNA was checked on a minigel (Section 2.10) to ensure complete digestion. If digestion was incomplete more enzyme was added and the reaction reincubated for another hour. If the digest failed to go to completion, the DNA was further purified by a phenol/chloroform extraction (Section 2.8 and 2.9). Once complete, the digestion was stopped by the addition of a one fifth volume of SDS loading dye (Section 2.4.2).

## 2.8 *Phenol/Chloroform Extraction*

DNA was purified by phenol/chloroform extraction to inactivate enzymes and other impurities that may interfere with further manipulations. The DNA solution was extracted with an equal volume of Tris-equilibrated phenol (Section 2.4.2) and chloroform, vortexed and spun for 5 min in a microfuge. The aqueous phase was then extracted with 1-2 volumes of chloroform as above. The DNA was precipitated with ethanol/isopropanol as described in Section 2.9.

## 2.9 *Precipitation of DNA*

A one tenth volume of 3 M sodium acetate (Section 2.4.2, either pH) and either 2.5 volumes of 95% ethanol or 0.6 volumes of isopropanol were added to the DNA solution. The solution was mixed by inversion and incubated on ice for a minimum of 15 min. The DNA was pelleted by centrifugation in a microfuge for at least 10 min. The DNA pellet was washed with 70% ethanol and dried at 37°C before resuspending in either $H_2O$ or TE (10/0.1, section 2.4.2).

## 2.10 *Agarose Gel Electrophoresis*

### 2.10.1 *Agarose Gels*

Agarose in either TBE or TAE buffers (Section 2.4.2) at the concentrations best suited for the desired separation described below, was melted by microwaving. The agarose was allowed to equilibrate to 50°C before pouring the gel. All gels were run in TBE buffer except when extracting DNA from a gel, when TAE buffer was used with a low melting point agarose.

| Grade of agarose | % (w/v) of agarose | Size range of separation |
|---|---|---|
| Molecular Biology (BRL) | 0.7 | 2 kb - 25 kb |
| Molecular Biology (BRL) | 1.0 | 0.5 kb - 10 kb |
| Molecular Biology (BRL) | 1.2 | 0.4 kb - 5 kb |
| NuSieve (FMC) | 2.0 | 0.3 kb - 1.5 kb |
| NuSieve (FMC) | 3.0 | 0.1 kb - 1 kb |
| SeaPlaque (FMC) | 0.7 | 2 kb - 25 kb |
| SeaPlaque (FMC) | 1.0 | 0.5 kb - 10 kb |

### 2.10.2 *Mini Gel Electrophoresis*

Horizontal agarose mini gels were run at 60-100 volts in TBE or TAE buffer (Section 2.4.2). Gels were run until the SDS loading dye (Section 2.4.2) had migrated approximately $^3/_4$ of the way down the gel.

### 2.10.3 *Overnight Gels*

Overnight gels were run overnight on a Biorad DNA Sub-Cell (150 x 200 mm gel bed) at 30 volts. Gels were run until the SDS loading dye (Section 2.4.2) had migrated at least 150 mm from the wells.

### 2.10.4 *Staining and Photographing Gels*

After electrophoresis the DNA was stained in an ethidium bromide bath (1 µg/ml of ethidium bromide in MilliQ $H_2O$) for 5 - 10 min for a mini gel, or 20 - 30 min for an overnight gel, and then destained in $H_2O$ for an equal time. The bands were visualised on a UV transilluminator and photographed with either Polaroid 667 film or an Alpha Innotech gel documentation system.

### 2.10.5 *Sizing DNA Bands*

DNA fragment sizes (kb) were determined after agarose gel electrophoresis by measuring the distance (mm) the fragment had migrated from the well. The size (in kb) was calculated by interpolation from a plot of known bands of a size standard (either λ *Hin*dIII-digested DNA or the 1 kb ladder, BRL) run on the same gel. Fragment sizes were also determined using the Alpha Innotech gel documentation system software.

## 2.11 *Southern Blotting and Hybridisation*

### 2.11.1 *Southern (Capillary) Blotting*

The blotting method used was based on that of Southern (1975). DNA to be transferred onto a Nylon membrane was separated by overnight gel electrophoresis and photographed

as described in section 2.10. The gel was gently agitated in the following series of solutions (section 2.4.2): Blotting solution 1 for 15 min, blotting solution 2 for 30 min and blotting solution 3 for at least 1 h. The gel was finally washed for 2 min in 2 x SSC (section 2.4.2) and assembled onto a blotting stand as follows. Two sheets of 3MM paper (Whatman) were used as wicks soaking in 20 x SSC (section 2.4.2), the sheets were covered with a film of gladwrap with a hole cut slightly smaller than the gel. The treated gel was placed on the stand, and a piece of Hybond N membrane (Amersham) that was cut slightly larger than the gel, was placed on top. Two pieces of 3 MM paper soaked in 2 x SSC (cut slightly smaller than the gel) were placed on top of the membrane, followed by two dry pieces then a stack of paper towels with a weight on top to keep the pile flat. After overnight transfer of DNA from gel to membrane, the membrane was removed, washed for 5 min in 2 x SSC and baked for 2 h at 80°C under vacuum.

### 2.11.2 *Radio-Labelled DNA Probe*

DNA was radio-labelled with [$\alpha$-$^{32}$P]dCTP (3 000 Ci/mmole, Amersham) using a High Prime Kit (Boehringer). After the labelling reaction the unincorporated label was removed by a Probe Quant spin column (Pharmacia). The labelled probe was boiled before addition to the pre-hybridised blot.

### 2.11.3 *Hybridisation of Radio-Labelled DNA Probe*

All hybridisations were performed in roller bottles in a Bachofer hybridisation oven. Membranes were pre-hybridised with 20 - 30 ml of 10 x Denhardts solution (section 2.4.2) for 2 h prior to the addition of the labelled probe. The filter was hybridised to the labelled probe where the 10 x Denhardts solution had been reduced to a maximum volume of 10 ml. Hybridisation occurred overnight at 65°C. Following hybridisation the filters were washed three times in 2 x SSC (section 2.4.2) for 15 min at room temperature before exposing to X-ray film (Fuji).

### 2.11.4 *DIG-Labelled DNA Probe*

DNA labelling with Digoxygenin (DIG)-11-dUTP was performed using a DIG-labelling kit (Boehringer). The concentration of the probe was determined by comparing serial dilutions of DIG-labelled sample to a labelled control supplied in the kit by spotting on a positively charged membrane (Boehringer).

### 2.11.5 *Hybridisation and Detection of DIG-Labelled DNA Probe*

Membranes to be hybridised to a DIG-labelled probe were pre-hybridised in DIG hybridisation buffer (Section 2.4.2) for 2 h at 65°C. The boiled DIG-labelled probe was added to 10 ml of DIG hybridisation buffer and incubated at 65°C overnight. After hybridisation the washes were as follows: twice in 2 x SSC (Section 2.4.2) with 0.1% SDS for 5 min; twice in 0.1 x SSC with 0.1% SDS for 15 min; 1 min in DIG buffer 1 (Section 2.4.2); 1 h in DIG buffer 2 (Section 2.4.2); 30 min in DIG buffer 2 with anti-Digoxygenin-AP (Boehringer); twice in DIG buffer 1 with 0.3% (v/v) Tween 20 (Sigma) for 15 min, then 2 min in DIG buffer 3 (Section 2.4.2). The membrane was incubated with lumigen PPD or CSPD (both Boehringer) for 5 min, placed between 2 acetate sheets, incubated for 15 min at 37°C and exposed to X-ray film (Fuji) for 15 min.

### 2.11.6 *Stripping Southern Blots*

A boiling solution of 0.1% (w/v) SDS was poured over the filter to be stripped and gently shaken until the solution had cooled to room temperature. This was repeated and the filter was checked by overnight exposure to X-ray film (Fuji).

## 2.12 *Library screening*

### 2.12.1 *Penicillium paxilli Library*

A λGEM-11 *P. paxilli* genomic library was made as described by Itoh and Scott (1994). The library was titred using the *E. coli* strain KW251 (grown overnight in LB broth, Section 2.3.4, supplemented with 10 mM $MgSO_4.7H_2O$ and 0.2 % (w/v) maltose) as a host, using serial dilutions as described below (Section 2.12.2).

### 2.12.2 *Library Plating*

Phage (100 µl) diluted to 2 x $10^4$ pfu/ml in SM buffer (section 2.4.2) was mixed with 100 µl of KW251 and incubated at 37°C for 30 min. The phage/KW251 mixture was added to 3 ml of top agarose (section 2.3.7), equilibrated to 50°C and poured onto LB plates. The

plates were incubated at 37°C for 6 to 8 hours until small plaques were visible. The plates were stored at 4°C.

### 2.12.3 *Filter Lifts and Hybridisation*

Filters (Hybond $N^+$, Amersham) were marked asymmetrically, placed on the phage plates from above (section 2.12.2) for 1 min and the marks transferred to the plates for later alignment. The filters were placed (DNA side up) on filter papers moistened with blotting solution 2 (section 2.4.2) for 2 min, blotting solution 3 (section 2.4.2) for 5 min and finally for 2 min on 2 x SSC (section 2.4.2). The filters were air dried and baked according to Section 2.11.1. To eliminate false positives, two lifts were taken from each plate where the second lift was left on the plate for 90 sec. Hybridisation was performed as described in section 2.11.3 or 2.11.5. Positive plaques were picked using a cut off blue tip, where there was alignment of the plate to the X-ray film and suspended in 0.5 ml of SM buffer (Section 2.4.2) with 50 µl of chloroform.

### 2.12.4 *Second and Third Round Screening*

Plaques picked for second and third round screening were plated as in Sections 2.12.2 and 2.12.3 with a lower titre per plate (between 100 - 300 plaques/plate). DNA was finally extracted as in Section 2.5.6.

## 2.13 *Subcloning*

---

### 2.13.1 *DNA Extraction from Agarose Gels*

DNA was recovered from SeaPlaque agarose gels (FMC) prepared in TAE buffer (Section 2.4.2) by phenol freeze extraction (Thuring *et al.*, 1975). After separation on a mini gel (Section 2.10.2) the DNA was stained using ethidium bromide and the band of interest excised from the gel under long wavelength UV light. The agarose was melted at 65°C, mixed with an equal volume of Tris-equilibrated phenol (Section 2.4.2) and frozen at -20°C for at least 2 h. The sample was centrifuged for 10 min, the aqueous phase recovered and extracted with phenol/chloroform (Section 2.8) and the DNA precipitated with ethanol (Section 2.9). The concentration of the excised fragment was determined by comparison with known standards on an agarose gel (Section 2.6.1).

### 2.13.2 *CAP-Treatment of Vector DNA*

Restriction enzyme digested vector DNA (5 µg, Section 2.7) was incubated with 0.1 unit of calf alkaline phosphatase (CAP, Boehringer) at 37°C for 30 min. $Na_2EDTA$ (final concentration of 5 mM), SDS (final concentration of 0.5% (w/v)) and proteinase K (final concentration of 50 µg/ml, Boehringer) were added, and the mixture was incubated at 56°C for 30 min. The DNA was purified by phenol/chloroform extraction (Section 2.8) and precipitated with ethanol (Section 2.9). The DNA was diluted to 10-20 ng/µl in MilliQ $H_2O$.

### 2.13.3 *Ligations*

Ligation was performed in a 20 µl reaction volume as follows: 2 µl of ligation buffer (New England Biolab), 20 ng of vector (Section 2.13.2), 0.1 µl of T4 DNA ligase (40 units, New England Biolab) and a 2-3 x molar excess of insert:vector. The ligation mixture was incubated overnight at 4°C To check ligation had occurred, two 2 µl samples of the ligation mixture were removed one before adding ligase and the other after ligation, mixed with SDS loading dye (Section 2.4.2) and separated on a minigel (Section 2.10.2).

### 2.13.4 *Plasmid Rescue*

Genomic DNA (50 ng) was digested with the appropriate restriction enzyme (Section 2.7), purified (Sections 2.8 and 2.9) and self-ligated using T4 DNA ligase (New England Biolab) in a 20 µl reaction volume, as described in Section 2.13.3.

### 2.13.5 *Preparation of Electrocompetent Cells*

One litre of LB (section 2.3.4) was inoculated with XL-1 ($^1/_{100}$, Section 2.1) and grown at 37°C with vigorous shaking to mid-log phase ($A_{600}$ 0.5 - 1.0, approximately 3 hours). The cells were chilled on ice for 20 min and harvested by centrifugation at 4 000 *g* for 10 min (4°C). The cells were resuspended twice in ice-cold water (1 litre, then 500 ml, spinning as above) followed by resuspension in ice-cold 10 % (v/v) glycerol (20 ml then 4 ml). The cells were stored at -70°C in 200 µl aliquots.

### 2.13.6 *Electroporation*

Electrocompetent *E. coli* cells (40 µl) were mixed with 2 µl of ligation mix (Sections 2.13.3 and 2.13.4) and incubated on ice for 1 min. The cells were placed in the base of an ice-cold 0.2 cm cuvette (BioRad), then pulsed at the following settings in a BioRad gene pulser: 25 µF, 2.5 kV and 200 ohms. When pulsing was complete, the cells were immediately resuspended in 1 ml of SOC medium (section 2.3.6) and incubated at 37°C for 1 h. The cells were plated at suitable dilutions onto selective LB plates (section 2.3.4). A positive control (20 ng of uncut pUC118) and a negative control (cells only) were always included in each experiment.

## 2.14 *DNA sequencing*

### 2.14.1 *Sequenase*

The Sequenase kit (USBC) was used for sequencing very pure DNA (Section 2.5.5). CsCl-purified DNA (5 µg) was denatured using 2 µl of a freshly prepared solution of 2 M NaOH, 2 mM $Na_2EDTA$ in a volume of 20 µl and incubated at room temperature for 5 min. The DNA was placed on ice and Tris-HCl (8 µl of 1 M, pH 8.0) and sodium acetate (3 µl of 3 M, pH 4.5, Section 2.4.2) were added. The solution was mixed well and 75 µl of ice-cold 95 % ethanol added. The DNA was pelleted by centrifugation for 10 min and the pellet washed with 200 µl of 70 % ethanol and dried. The DNA was resuspended in 7 µl of MilliQ $H_2O$, 1 µl of appropriate sequencing primer (1 pmol) and 2 µl of 5 x sequenase sequencing buffer added, and the mixture incubated at 37°C for 30 min. After incubation, the following reagents were added: 1 µl of 0.1 M dithiothreitol, 2 µl of labelling mix (diluted 5 x in MilliQ water), 0.5 µl of [$\alpha$-$^{35}$S]dATP (>1000 Ci/mmole, Amersham) and 2.0 µl of Sequenase (diluted 8 x in the supplied enzyme-dilution buffer). The DNA was incubated for another 5 min at room temperature. The four termination mixes (2.5 µl) were then incubated with 3.5 µl of the above labelling mix for 5 min at 37°C. The reaction was stopped by the addition of 4 µl of stop solution and stored at -20°C.

### 2.14.2 *Amplicycle Sequencing*

DNA that was prepared by a mini alkaline lysis method (section 2.5.4) was sequenced using the Amplicycle sequencing kit (Perkin-Elmer). A 30 µl cocktail with 1 µg of DNA,

4 µl of cycling mix, 1 µl of appropriate primer (1 pmol) and 1 µl of [$\alpha$-$^{33}$P]dCTP (~2500 Ci/mmole, Amersham) was mixed with 2 µl of termination mixes and placed in a Corbett thermocycler set to 94°C. The following programme was used: 94°C for 1 min followed by 25 cycles of 94°C for 30 sec, 50°C for 30 sec and 72°C for 60 sec. Stop solution (4 µl) was added after the last cycle.

### 2.14.3 *Automatic Sequencing*

DNA extracted by CsCl and alkaline lysis (Section 2.5) were sequenced using an ABI 373 and 377 automatic sequencer with Prism (ABI) dye terminator chemistry.

### 2.14.4 *Electrophoresis of Sequencing Reactions*

Radio-labelled sequencing reactions were separated by polyacrylamide gel electrophoresis. Sequencing gels were poured with 60 ml of acrylamide mix (Section 2.4.2) containing 36 µl of TEMED and 360 µl of 10 % (w/v) ammonium persulphate. Once set, the gels were run at 1500 V for 15 min in 1 x sequencing TBE (Section 2.4.2). The sequencing reactions were denatured at 80°C for 2 min prior to loading 3 µl onto the gel. The samples were run for either a short run of 2 h (until the first blue dye front ran off) or a long run of 6 h (when 3 dye fronts had run off). The gels were fixed in 10 % (v/v) acetic acid, 10 % (v/v) ethanol for 15 min, dried for 35 min under vacuum at 80°C, then exposed to X-ray film (Fuji) overnight.

### 2.14.5 *Sequence Analysis*

DNA sequence was analysed with the Wisconsin Package version 9.1, Genetics Computer Group (GCG), Madison, Wisc. The programs used were the Fragment assembly package, MAP and FRAMES. Blast searches (blastX) were preformed through Netscape Navigator at NCBI using the Brookhaven, SWISSPROT and GenBank databases.

## 2.15 PCR

### 2.15.1 *Primers (Oligonucleotides)*

Each oligonucleotide was synthesised by either "Oligo etc" or "BRL". Each primer was rehydrated to a final concentration of 100 mM. For PCR analysis, primers were used at a stock concentration of 10 mM. For sequencing reactions primers were used at a 1 mM

stock concentration. The sequence of all primers used in this thesis are described in Table 2.2.

### 2.15.2 *PCR for Site Specific Integration*

All PCR reactions were carried out in a final volume of 25 µl containing 1 x *Taq* polymerase buffer (Boehringer), 50 µM of each dNTP, 200 nM of each primer, 2 units of *Taq* DNA polymerase (Boehringer) and 5 ng of genomic DNA. The temperature cycles were performed in a Corbett thermocycler using the programmes described below.

| Primer combination | PCR programme |
|---|---|
| Pax12/Hph1 | 94°C, 2 min; 30 cycles of 94°C, 45 sec; 56°C, 45 sec; 72°C, 90 sec, finishing with 72°C for 5 min. |
| Pax7/Pax8 | 94°C, 2 min; 30 cycles of 94°C, 45 sec; 56°C, 45 sec; 72°C, 90 sec, finishing with 72°C for 5 min. |
| Pax13/Hph1 | 94°C, 2 min; 30 cycles of 94°C, 45 sec; 60°C, 45 sec; 72°C, 2.30 min, finishing with 72°C for 5 min. |
| Pax14/pUC reverse | 94°C, 2 min; 30 cycles of 94°C, 45 sec; 55°C, 45 sec; 72°C, 2.30 min, finishing with 72°C for 5 min. |

The PCR products were identified by separation on agarose gels (Section 2.10) in 1 x TBE buffer (Section 2.4.2).

### 2.15.3 *Random Amplification of Polymorphic DNA (RAPD)*

The amplification conditions used are based on those of Williams *et al.* (1990) using modifications as described by Crowhurst *et al.* (1991). Amplification reactions were carried out in a final volume of 25 µl containing 1 x *Taq* polymerase buffer (Boehringer), 100 µM of each dNTP, 200 nM of each primer (GT02, RC05 or RC09, Section 2.15.1), 1 unit of *Taq* DNA polymerase (Boehringer) and 25 ng of genomic DNA. The PCR programme used, 94°C, 3 min; 40 cycles of 94°C, 1 min; 35°C, 1 min; 72°C, 3 min. The

*Table 2.2* *Primer sequences.*

| Primer name | Primer sequence | Application |
|---|---|---|
| GT02 | 5'-TGGTGGGTCC-3' | RAPD |
| hph1 | 5'-GCTTGATTGGCTGCTATAAC-3' | PCR |
| pax1 | 5'-GTCACCCTTCCGAACAGGAG-3' | Seq |
| pax2 | 5'-TGCCGGAATTCATAATCCCG-3' | Seq |
| pax3 | 5'-ATCGAAGCCTGCTGCTACGG-3' | Seq |
| pax4 | 5'-ATCTGCAGGTTCAATCTCGG-3' | Seq |
| pax5 | 5'-TTGCGAGAGAGACTCCGCTG-3' | Seq |
| pax6 | 5'-CGGAGATGTACTAGGAATAC-3' | Seq |
| pax7 | 5'-CAATGGCATCATCGGTGTCC-3' | PCR |
| pax8 | 5'-ACATCCATAACAGTCGACTC-3' | PCR |
| pax9 | 5'-ATCCATACCAGGACAGGTAG-3' | Seq |
| pax10 | 5'-ATCTTCGTGGCAATGAACTA-3' | Seq |
| pax11 | 5'-AGTTATCCGCAAGAAGTGCA-3' | Seq |
| pax12 | 5'-CCTGGAATTCTGTCGTATAA-3' | PCR/Seq |
| pax13 | 5'-GTGGATTCTGCGATGACTTG-3' | PCR |
| pax14 | 5'-GGTCGTCGGAGAATACGGAG-3' | PCR |
| pax19 | 5'-TCACGGTTGTACCCGCAAAG-3' | Seq |
| pUCap | 5'-GTCAATACGGGATAATACCG-3' | IPCR |
| pUC forward | 5'-GTTTTCCCAGTCACGAC-3' | Seq |
| pUC reverse | 5'-CAGGAAACAGCTATGAC-3' | PCR/IPCR/Seq |
| RC05 | 5'-AGGAGATACC-3' | RAPD |
| RC09 | 5'-GATAACGCAC-3' | RAPD |

amplified products were visualised by overnight agarose gel (1% w/v) electrophoresis (Section 2.10). Primers were kindly provided by Ross Crowhurst, HortResearch, Auckland, New Zealand.

### 2.15.4 *Inverse PCR (IPCR)*

Restriction enzyme digested DNA (50 ng) was ligated (Section 2.13.4) and underwent PCR with the pUC reverse and pUCap primers using the same reaction conditions as in Section 2.15.2. The following PCR programme was performed in a Corbett thermocycler: 94°C, 2 min; 30 cycles of 94°C, 30 sec; 55°C, 30 sec; 72°C, 90 sec; finishing with 72°C, 5 min.

## 2.16 *Fungal Protoplast Protocols*

### 2.16.1 *Fungal Protoplasts*

Protoplasts of PN2013 were prepared using modifications of the method described by Yelton *et al.* (1984), Murray *et al.* (1992) and Itoh *et al.* (1994). Fungal cultures (inoculated with 5 x $10^6$ spores) were grown in 35 ml of PD broth at 22°C with vigorous shaking (200 rpm) for 40 h. The mycelial pellets were washed 3 times with sterile MilliQ $H_2O$ followed by a wash in OM buffer (Section 2.4.2). The washed mycelia were mixed with 15 ml of sterile OM buffer containing 10 mg/ml of Novozyme 234 (Novo Industri A/S). This mixture was shaken gently (80 rpm) for 2-3 h at 28°C. The digested hyphae were filtered through mira-cloth (Calbiochem) and the filtrate containing protoplasts overlayed with 2 ml of ST buffer (Section 2.4.2). The protoplasts were banded at the interface by centrifuging at 3 000 *g* for 5 min. After removing from the interface the protoplasts were washed 3 times with 5 ml of STC buffer (Section 2.4.2) by centrifuging at 7 700 *g*. Protoplasts made for transforming (section 2.16.2) were resuspended in STC buffer to give a final concentration of 1.25 x $10^8$ per ml, while those used for CHEF analysis (section 2.16.3) were resuspended in GMB (Section 2.4) buffer to a final concentration of at least 5 x $10^8$ protoplasts/ml.

### 2.16.2 *Fungal Transformation*

Protoplasts, prepared as described above (Section 2.16.1) were transformed by the method of Vollmer and Yanofsky (1986), as modified by Itoh *et al.* (1994). Each transformation used 5 µg of DNA with 80 µl of protoplasts, 5 µl heparin (5 mg/ml in STC buffer), 2 µl spermidine (50 mM) and 20 µl PEG buffer (Section 2.4.2). After incubation on ice for 30 min, 900 µl of PEG buffer (Section 2.4.2) was added to the protoplasts and incubated at room temperature for 15 min. A 100 µl aliquot of the transformation mix was plated onto complete media plates (Section 2.3.2) with a 5 ml complete media top agar (Section 2.3.3) overlay. The plates were incubated at 22°C overnight then overlayed as above with hygromycin B (Boehringer) to a final plate concentration of 200 µg/ml. The resulting transformants were subsequently maintained on PD media with a hygromycin B concentration of 20 µg/ml.

### 2.16.3 *Plug Preparation*

Protoplasts, prepared as described above (section 2.16.1), were mixed with an equal volume of LMP in GMB (Section 2.4.2) and allowed to set in plug molds (BioRad) for 10 - 15 min. The plugs were then incubated at 50°C for 18 h in SE buffer (Section 2.4.2) followed by a 24 h incubation at 50°C in 10 ml of 10 x ET (Section 2.4.2) with 20 mg of proteinase K (Boehringer) and 100 mg sodium lauroylsarcosine (Sigma). The plugs were finally washed four times with 1 x ET (Section 2.4.2) buffer and stored at 4°C.

### 2.16.4 *Karyotype Analysis*

Chromosomal DNA of *P. paxilli* was separated by contour-clamped homogenous electric field (CHEF)-gel electrophoresis as described in Itoh *et al.* (1994). A BioRad CHEF DRII was used for all chromosome analysis. Plugs were initially checked for quality and concentration on a 0.6 % (w/v) chromosomal grade agarose (BioRad) in 0.5 x TBE (section 2.4.2) using the following programme: 100 - 1000 sec switching, 100 volts for 19 h at 10°C. For a chromosome separation two programs were used. The program of Itoh *et al.* (1994) is as follows: 0.6% (w/v) chromosomal grade agarose in 0.5 x TBE at 60 volts with the following switching times 500 sec for 25 h, 1000 sec for 30 h and 1500 sec for 50 h and 2500 sec for 32 h. The modified program of Geiser *et al.* (1996) is as follows: 0.6% (w/v) chromosomal grade agarose in 0.5 x TBE at 47 volts with the following switching times 3000 sec for 73 h, 2700 sec for 18 h and 2200 sec for 73 h.

## 2.17 *Paxilline Assays*

### 2.17.1 *Monoclonal-Antibody ELISA*

Paxilline analysis of PN2013 and resulting transformants was carried out at AgResearch, Ruakura, Hamilton, NZ. Approximately $10^6$ spores were inoculated into 3 ml of CDYE (section 2.3.1), then grown for 7-10 days with vigorous shaking (200 rpm) at 28°C. Culture filtrates were analysed for paxilline by a competitive ELISA using a group-specific monoclonal antibody, MYCO-03/01 (Garthwaite *et al.*, 1993). This antibody has a dynamic working range of 2 - 1 000 ng/ml in the assay. Cultures that proved negative were grown up a second time in 25 ml and analysed by both ELISA and HPLC (section 2.17.2).

### 2.17.2 *HPLC Analysis for Paxilline*

Cultures (5 x $10^6$ spores) were inoculated in 25 ml of CDYE (section 2.3.1) and allowed to grow for 7 - 10 days at 28°C. The sample was freeze dried, ground under liquid $N_2$ and extracted in 15 ml of 2:1 chloroform:methanol. The samples were shaken for 1-2 hours at 600 rpm on a Vibrofix shaking platform then allowed to settle for 30 min. A 1.5 ml sample was removed, solvent evaporated under a stream of $N_2$, then redissolved in 1 ml of dichloromethane (DCM). Before HPLC analysis, the samples were purified using Sep-Pak plus cartridges (Waters). The cartridges were pre-wet with 3 ml of DCM, the sample added, then washed with 1 ml of DCM:acetonitrile (80:20). The sample was eluted with 3 ml of DCM:acetonitrile (20:80) and 10 µl was analysed by normal phase HPLC using a Zorbax (Rockland Technologies) silica gel column (4.6 mm x 25 cm, 5 µm) and a mobile phase of DCM:acetonitrile (80:20) containing 0.4% methanol (1.8 ml/min). Eluted products were detected with a Hewlett-Packard 1040M diode array UV detector set at 230 nm. Paxilline was quantitated by absorbance at 230nm. The sensitivity of this assay is of the order of 5 - 10 ng paxilline per ml.

### 2.17.3 *TLC Analysis for Paxilline*

Samples analysed for paxilline by TLC were extracted as above (Section 2.17.2). However, after drying a 1.5 ml sample, the sample was then resuspended in 50 µl of chloroform. 15 - 20 µl sample spots were separated on SIL G (Macherey-Nagel) TLC

plates using 10:1 chloroform:acetone as the running solvent. For better separation the plate was run for two solvent fronts. Paxilline was identified by spraying the plate with alcoholic Ehrlich's reagent (2% (w/v) p-dimethylaminobenzaldehyde in 12% (v/v) HCl and 50% (v/v) ethanol) and heating the plate to 120°C for approximately 10 min.

# *Chapter Three*
# *Results*

## 3.1 *Mapping Genomic Sequences Flanking the Site of pAN7-1 Integration*

Previous results by Itoh (described in the Introduction) show that the integration of pAN7-1 in YI-20 is at a single site but that the plasmid is present as a tandem repeat. Two possible *Bam*HI fragments (13.6 kb and 12.1 kb) containing the flanking sequence can be seen in a *Bam*HI digest of YI-20 DNA, when probed with pAN7-1 (Fig 1.5 and 3.2). It appears from these Southern results that the flanking *Bam*HI sites of pAN7-1 have been lost during integration but retained within the tandem repeat as can be seen by the presence of a 6.8 kb hybridising fragment. As no other potential flanking bands were detected in other digests, smaller fragments of pAN7-1 were used as probes. Probe A, a 0.8 kb *Bam*HI/*Eco*RI fragment (Fig. 3.1A), and probe B, a 0.8 kb *Hin*dIII/*Bam*HI fragment (Fig. 3.1A) would be the left and right ends of pAN7-1 when linearised with *Bam*HI (Fig. 3.1A). On this basis, and the fact that the hybridisation patterns would be simpler than using the complete pAN7-1 probe, these two fragments were used to determine a map of the surrounding genomic flanking sequence.

### 3.1.1 *Mapping the Left-Hand Side of the pAN7-1 Integration*

Initially, probe A ([$\alpha$-$^{32}$P]dCTP-labelled 0.8 kb *Bam*HI/*Eco*RI fragment, Fig. 3.1A) was used to probe various restriction enzyme digests of YI-20 (Section 2.11). The size of the hybridising fragments, due to the tandem repeat, were based on the fragment sizes of the available pAN7-1 sequence (accession number Z32698). These fragments are seen as strong hybridising bands in Figure 3.2 and are described in Table 3.1. Fainter bands relating to the pAN7-1 point of integration in the YI-20 genome (Fig. 3.2), are as follows; 3.6 kb *Sst*I fragment (lane 1), 1.8 kb *Eco*RI fragment (lane 3), also present in the following double digests *Eco*RI/*Sst*I, *Eco*RI/*Bam*HI (lanes 2 and 4), 12.1 and 13.6 kb *Bam*HI fragments (lane 5) and 6.8 kb *Hin*dIII/*Bam*HI (lane 6). Other faint bands, described in Table 3.1, are due to pAN7-1 carryover while isolating the 0.8 kb *Bam*HI/*Eco*RI fragment (Section 2.13). From these data a map was determined using the fact that the initial integrating plasmid was *Bam*HI linearised (Fig. 3.1B). Although the *Bam*HI site adjacent to the genomic flanking sequence has been destroyed, since probe A detects the presence of single copy bands (Fig. 3.2), this indicates that very little of the adjacent pAN7-1 sequence was lost during its integration. On this basis, a map (Fig. 3.1B) can be drawn by subtracting the expected size of pAN7-1 contained within the hybridising fragments. Therefore, the 3.6 kb *Sst*I fragment contains approximately 1.6

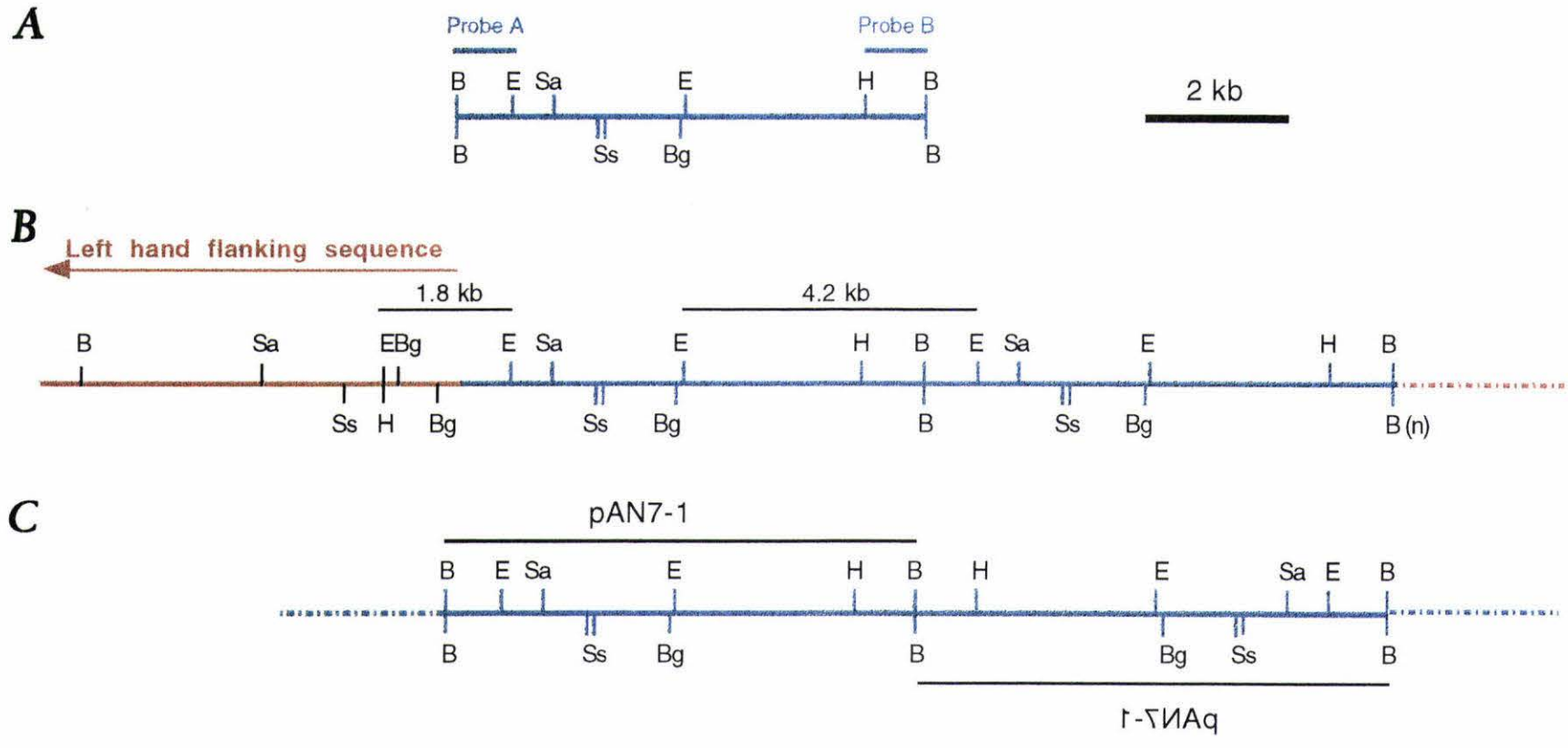



***Fig. 3.1*** *Restriction enzyme maps of pAN7-1 and the integration events in YI-20.*

***A*** Restriction enzyme map of *Bam*HI linearised pAN7-1.

***B*** Restriction enzyme map of the left hand integration event of pAN7-1 in YI-20.

***C*** Restriction enzyme map of an inverted (head to head) copy of pAN7-1.

Abbreviations: B, *Bam*HI. Bg, *Bgl*II. E, *Eco*RI. H, *Hin*dIII. Sa, *Sal*I. Ss, *Sst*I. Blue lines are pAN7-1. Red lines are YI-20 flanking sequence.

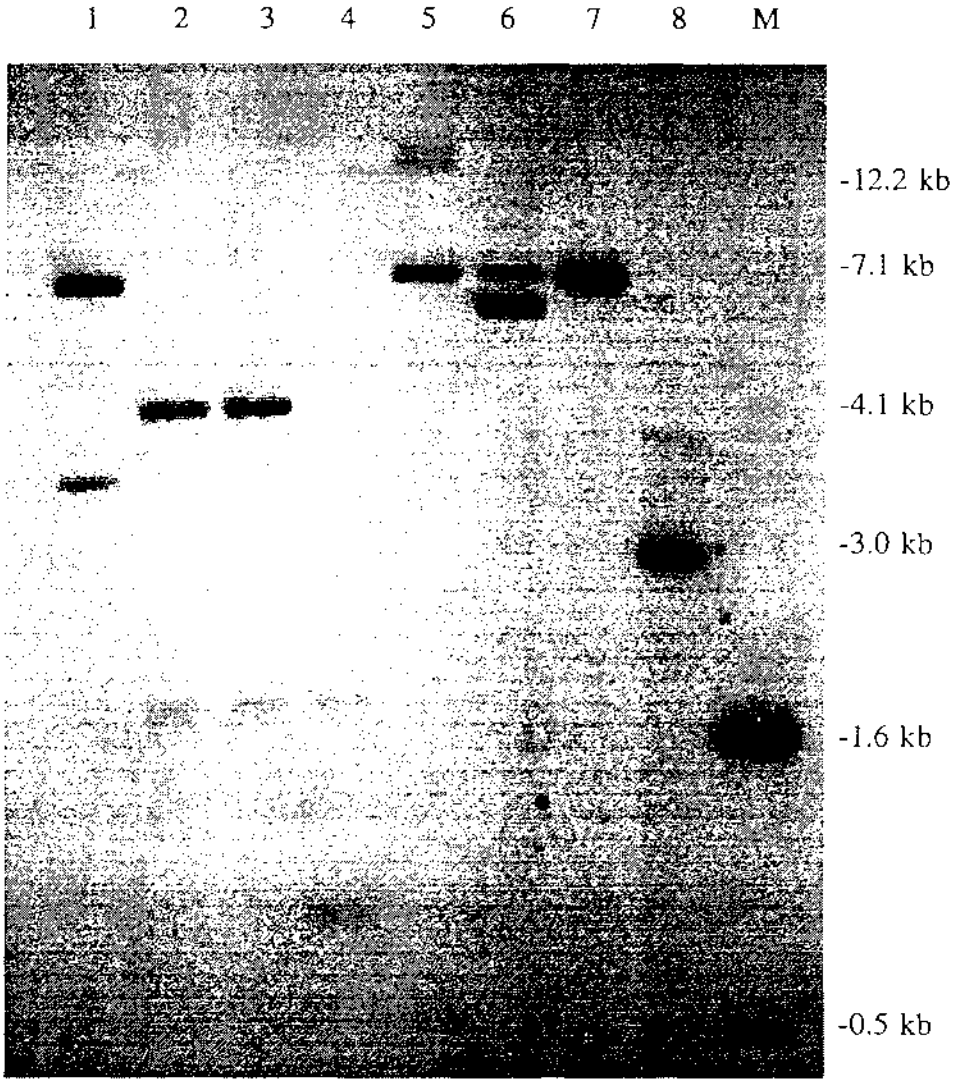



**Fig. 3.2** *Mapping of the YI-20 left-hand flanking sequence with respect to the pAN7-1 insertion.*

Southern analysis of 1 µg of YI-20 DNA digested with *Sst*I (lane 1), *Eco*RI/*Sst*I (lane 2), *Eco*RI (lane 3), *Bam*HI/*Eco*RI (lane 4), *Bam*HI (lane 5), *Hin*dIII/*Bam*HI (lane 6), *Hin*dIII (lane7), *Hin*dIII/*Sst*I (lane 8), hybridised with the 0.8 kb *Bam*HI/*Eco*RI fragment (probe A) of pAN7-1. Lane M, 1 kb Ladder.

**Table 3.1** *Size of bands that hybridise to the 0.8 kb EcoRI/BamHI fragment from pAN7-1.*

| Enzyme Digest | Band containing flanking sequence | Band containing pAN7-1 tandem repeat | Band containing other pAN7-1 fragments[a] |
|---|---|---|---|
| *Sst*I | 3.6 kb | 6.7 kb | |
| *Eco*RI/*Sst*I | 1.8 kb | 4.2 kb | |
| *Eco*RI | 1.8 kb | 4.2 kb | 2.5 kb |
| *Bam*HI/*Eco*RI | 1.8 kb | 0.8 kb | 2.5 kb |
| *Bam*HI | 12.1 kb | 6.8 kb | 13.6 kb |
| *Hin*dIII/*Bam*HI | 6.8 kb | 6.0 kb | |
| *Hin*dIII | 6.8 kb[b] | 6.8 kb | |
| *Hin*dIII/*Sst*I | 2.8 kb[b] | 2.8 kb | 3.8 kb |

[a]Other pAN7-1 bands are due to pAN7-1 contamination of the probe.
[b]Hybridises to a band of the same size as the tandem repeat.

kb of genomic flanking sequence (Fig. 3.1B). The 1.8 kb *Eco*RI fragment contains 1.1 kb of flanking sequence that does not include either *Bam*HI or *Sst*I sites (Fig 3.1B). The expected *Hin*dIII fragment (Fig. 3.2, lane 7) is hidden by hybridisation of the tandem repeat. However, the 6.8 kb fragment seen in the *Hin*dIII/*Bam*HI digest (Fig. 3.2, lane 6) confirms this size. This fragment must contain at least 6.0 kb of pAN7-1 with the remaining 0.8 kb being flanking sequence (Fig. 3.1B). This placement of the *Hin*dIII site is further confirmed by the *Hin*dIII/*Sst*I digest as the expected single copy band would be hidden by the strong hybridising 2.8 kb tandem repeat band (Fig. 3.2, lane 8). At this stage the *Bam*HI site was unable to be placed because there are two potential hybridising fragments (13.6 kb and 12.1 kb, Fig. 3.2). Data shown later confirms that the smaller fragment of 12.1 kb contains 5.4 kb of left-hand flanking sequence.

### 3.1.2 *An Inverted Copy of pAN7-1*

An assumption was made that a map of the right-hand side could be determined by hybridisation with the 0.8 kb *Hin*dIII/*Bam*HI fragment (probe B, Fig. 3.1A) from pAN7-1. However, hybridisation of a Southern blot with this fragment (Fig. 3.3) was not very informative as the fainter hybridising bands: (i) 9.3 kb *Sst*I fragment (lane 1), (ii) 6.8 kb *Eco*RI fragment (lane 3) also seen in the *Eco*RI/*Sst*I and *Eco*RI/*Bam*HI digests (lanes 2 and 4), (iii) 13.6 kb *Bam*HI fragment (lane 5), (iv) 1.6 kb *Hin*dIII fragment (lane 7) also seen in the *Hin*dIII/*Bam*HI and *Hin*dIII/*Sst*I digests (lane 6 and 8), are sizes predicted from an inverted copy of pAN7-1, where the *Bam*HI site was deleted between the two copies (Fig. 3.1C). Fragment sizes of such a pAN7-1 integration match those predicted from the actual sequence data (accession number Z32698) and this head-to-head integration must be placed within the tandem repeat. Other fragments hybridising to the pAN7-1 tandem repeat (strong hybridising bands) or due to plasmid carry-over (faint bands) are shown in Table 3.2.

To confirm the presence of an inverted copy of pAN7-1, YI-20 genomic DNA was digested with a range of restriction enzymes, and probed with either DIG-labelled pAN7-1 (Section 2.11) or [$\alpha$-$^{32}$P]dCTP-labelled 0.8 kb *Bam*HI/*Eco*RI fragment (probe A, Fig. 3.1A). Bands due to the inverted repeat would not hybridise to the smaller probe (probe A). When probed with pAN7-1 (Fig. 3.4A) the inverted repeat fragments were seen as less intense bands; 6.8 kb *Eco*RI fragment (lane 1), 9.3 kb *Sst*I fragment (lane 5), 7.1 kb *Bgl*II fragment (lane 6) also in the *Bgl*II/*Sst*I digest (lane 7) and the 10.7 kb *Sal*I fragment (lane 8) also in the *Sal*I/*Cla*I digest (lane 9). The remaining fragments are the result of hybridisation to the tandem repeat (seen as intense hybridising fragments) or to

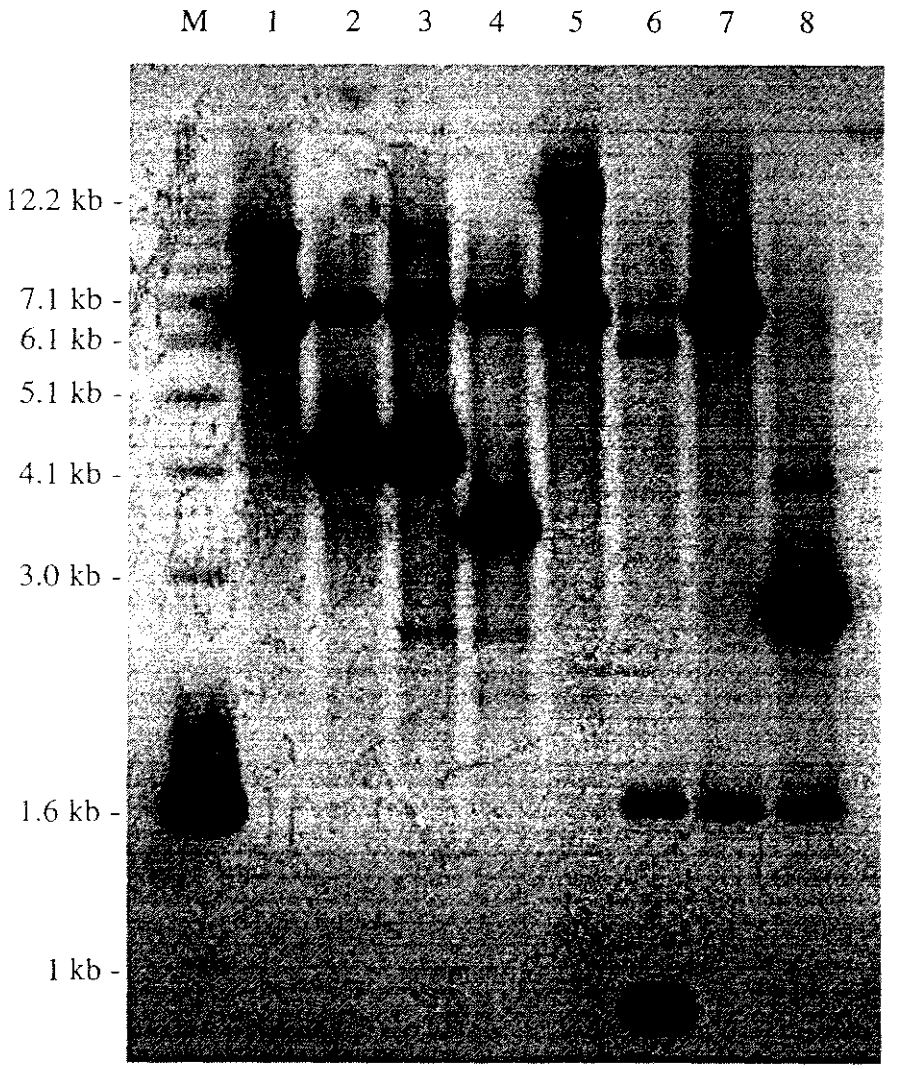



**Fig. 3.3** *Mapping of YI-20 showing the inverted repeat of pAN7-1.*

Southern analysis of 1 µg of YI-20 DNA digested with *Sst*I (lane 1), *Eco*RI/*Sst*I (lane 2), *Eco*RI (lane 3), *Bam*HI/*Eco*RI (lane 4), *Bam*HI (lane 5), *Hin*dIII/*Bam*HI (lane 6), *Hin*dIII (lane7), *Hin*dIII/*Sst*I (lane 8), hybridised with the 0.8 kb *Bam*HI/*Hin*dIII fragment (probe B) of pAN7-1. Lane M, 1 kb Ladder.

***Table 3.2*** *Size of bands that hybridise to the 0.8 kb HindIII/BamHI fragment from pAN7-1.*

| Enzyme Digest | Band containing inverted repeat | Band containing pAN7-1 tandem repeat | Band containing other pAN7-1 fragments[a] |
|---|---|---|---|
| *Sst*I | 9.3 kb | 6.7 kb | |
| *Eco*RI/*Sst*I | 6.8 kb | 4.2 kb | |
| *Eco*RI | 6.8 kb | 4.2 kb | 2.5 kb |
| *Bam*HI/*Eco*RI | 6.8kb | 3.4 kb | 2.5 kb |
| *Bam*HI | 13.6 kb | 6.8 kb | 12.1 kb |
| *Hin*dIII/*Bam*HI | 1.6 kb | 0.8 kb | 6.0, 6.8 kb |
| *Hin*dIII | 1.6 kb | 6.8 kb | |
| *Hin*dIII/*Sst*I | 1.6 kb | 2.7 kb | 3.8 kb |

[a]Other pAN7-1 bands are due to pAN7-1 contamination of the probe.

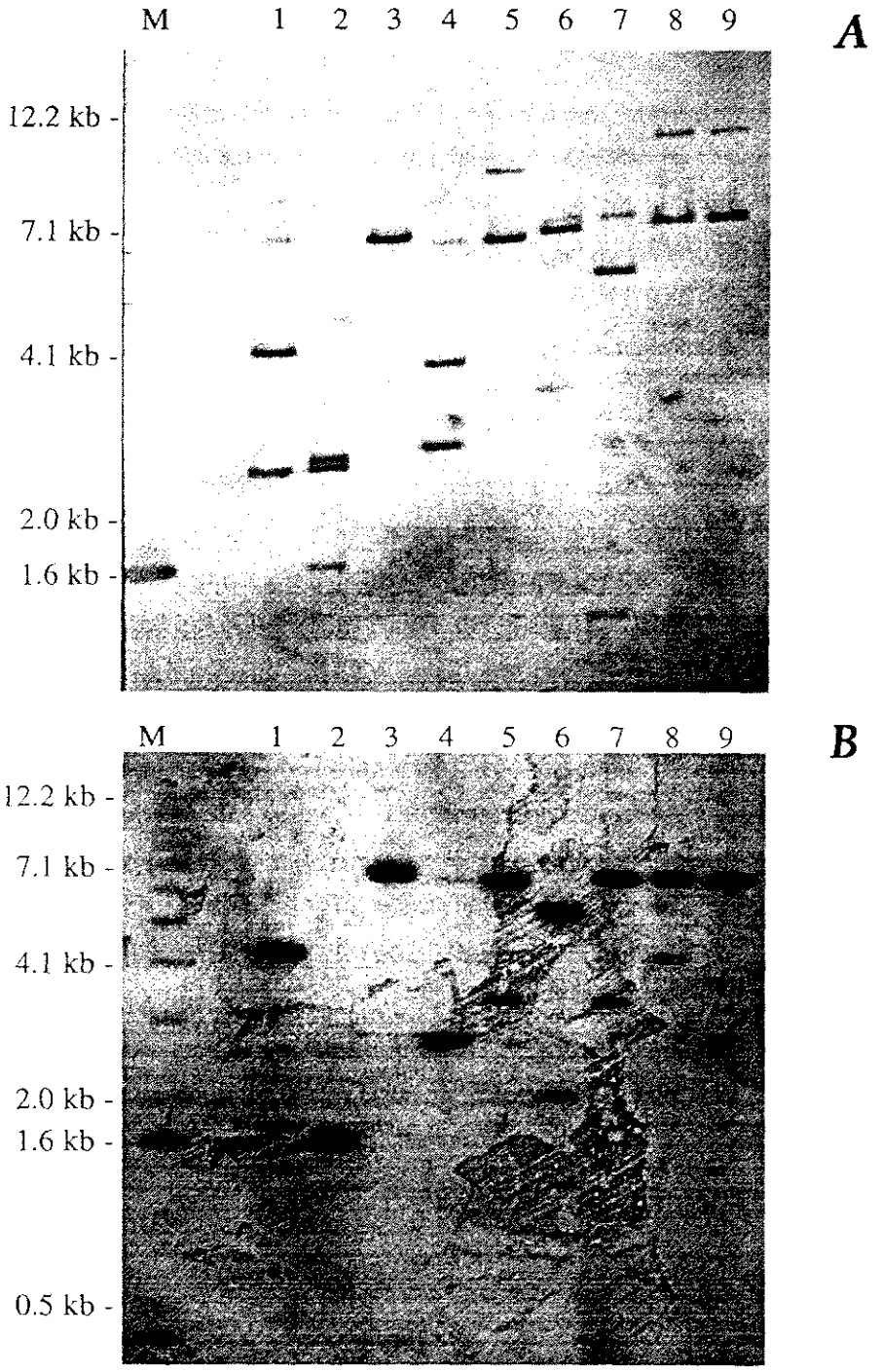



**Fig. 3.4** *Confirmation of the inverted copy of pAN7-1 in YI-20*

Southern analysis of 1 µg of YI-20 genomic DNA digested with:

***A*** *Eco*RI (lane 1), *Eco*RI/*Hin*dIII (lane 2), *Hin*dIII (lane 3), *Hin*dIII/*Sst*I (lane 4), *Sst*I (lane 5), *Bgl*II (lane 6), *Sst*I/*Bgl*II (lane 7), *Sal*I (lane 8), *Sal*I/*Cla*I (lane 9), hybridised with DIG-labelled pAN7-1.

***B*** *Eco*RI (lane 1), *Eco*RI/*Hin*dIII (lane 2), *Hin*dIII (lane 3), *Hin*dIII/*Sst*I (lane 4), *Sst*I (lane 5), *Sst*I/*Bgl*II (lane 6), *Bgl*II (lane 7), *Sal*I (lane 8), *Sal*I/*Cla*I (lane 9), hybridised with [α-$^{32}$P]dCTP-labelled 0.8kb *Bam*HI/*Eco*RI fragment (probe A) of pAN7-1.

the left-hand flanking region and are listed in Table 3.3. When the same digests are hybridised to [α-$^{32}$P]dCTP-labelled 0.8 kb *Bam*HI/*Eco*RI fragment (probe A, Fig. 3.1A), only bands relating to the tandem repeat and left-hand flanking region hybridise (Fig. 3.4B), and are listed in Table 3.3. These data allow placement of the *Bgl*II and *Sal*I sites on the left-hand flanking sequence (Fig. 3.1B). The 3.3 kb single copy *Bgl*II fragment has 0.3 kb of left-hand flanking sequence while the 4.2 kb *Sal*I fragment consists of 2.9 kb.

Further evidence for a pAN7-1 inverted repeat is the presence of a 13.6 kb *Bam*HI fragment seen in previous results (Fig. 3.2 and 3.3). A Southern blot of *Bam*HI-digested YI-20 genomic DNA (Section 2.11) shows the following hybridising fragments when probed with [α-$^{32}$P]dCTP-labelled pAN7-1 (Fig. 3.5, lanes 1 and 6): 6.8 kb tandem repeat band, 12.1 kb left-hand flanking band and 13.6 kb inverted repeat band. The 12.1 kb *Bam*HI fragment shows a shift in size when digested with enzymes (*Cla*I, *Eco*RV, *Kpn*I and *Sma*I) that do not cut within pAN7-1. In three of the four *Bam*HI double digests with these enzymes, a clear drop in the size of the 12.1 kb fragment can be seen (Fig. 3.5, lanes 2, 4 and 5). As expected the pAN7-1 tandem repeat (6.8 kb) and the inverted repeat (13.6 kb) remain intact. The 8.3 kb *Bam*HI/*Cla*I fragment (lane 5), 8.8 kb *Bam*HI/*Eco*RV fragment (lane 2) and the 10.1 kb *Bam*HI/*Kpn*I fragment (lane 4) contain 3.8 kb, 3.3 kb and 2 kb of left-hand flanking sequence, respectively. *Sma*I (lane 3) does not cut the 12.1 kb fragment. These data confirm that the 12.1 kb *Bam*HI fragment (lanes 1 and 6) contains the adjacent flanking sequence (5.4 kb), allowing placement of the *Bam*HI site (Fig. 3.1B).

## 3.2 *pAN7-1 Copy Number*

To determine the pAN7-1 copy number in YI-20, a Southern blot approach was used. An *Eco*RI digest of YI-20 DNA, when hybridised with the 0.8 kb *Bam*HI/*Eco*RI fragment (probe A, Fig. 3.1A), results in a 1.8 kb single-copy fragment, which corresponds to left-hand flanking sequence (Fig. 3.1B and 3.6) and a 4.2 kb multi-copy fragment, due to the tandem repeat (Fig. 3.1B and 3.6). The intensity of the multi-copy fragment (4.2 kb) was compared to that of the 1.8 kb fragment as the probe hybridises to both fragments equally. Therefore, the copy number of the two hybridising fragments is determined as 1 + ratio. Although there is a head-to-head copy of pAN7-1 in YI-20, this does not interfere with the above result as the probe does not hybridise to the 6.8 kb head-to-head *Eco*RI fragment (Fig. 3.1C). However, this fragment is observed because of hybridisation of a small amount of pAN7-1 carryover in the isolated probe. Using

**Table 3.3** *Size of bands that hybridise to pAN7-1 and probe A.*

| | pAN7-1 probe | | | |
|---|---|---|---|---|
| | | | Probe A | |
| Enzyme Digest | Band containing Inverted copy | pAN7-1 Tandem repeat | Left hand flanking sequence | pAN7-1 Tandem repeat |
| *Eco*RI | 6.8 kb | 2.5, 4.2 kb | 1.8 kb[a] | 4.2 kb |
| *Eco*RI/*Hin*dIII | 1.6 kb[b] | 1.6, 2.5, 2.7 kb | 1.6 kb[b] | 1.6 kb |
| *Hin*dIII | 1.6 kb[a] | 6.8 kb | 6.8 kb[b] | 6.8 kb |
| *Hin*dIII/*Sst*I | 1.6kb[a] | 2.7, 3.8 kb | 6.8 kb | 2.7 kb |
| *Sst*I | 9.3 kb | 6.7 kb | 3.3 kb | 6.7 kb |
| *Bgl*II | 7.1 kb | 6.8 kb | 3.3 kb | 6.8 kb |
| *Bgl*II/*Sst*I | 7.1 kb | 1.0, 5.4 kb | 2.0 kb | 5.4 kb |
| *Sal*I | 10.7 kb | 6.8 kb | 4.2 kb | 6.8 kb |
| *Sal*I/*Cla*I | 10.7 kb | 6.8 kb | 2.8 kb | 6.8 kb |

[a]Not seen in this exposure.

[b]Not visible as a faint band because it is the same size as the tandem repeat.

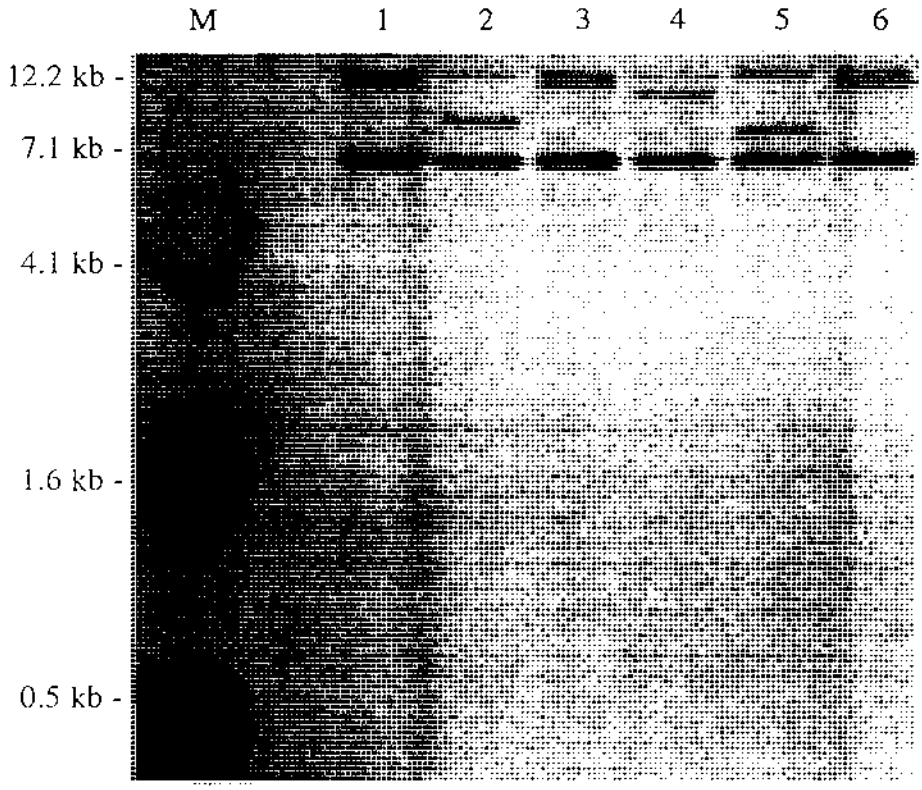



***Fig. 3.5*** *Southern analysis confirming that the 13.4 kb BamHI hybridising band is due to a pAN7-1 inverted copy.*

Southern analysis of 1 μg of YI-20 genomic DNA digested with *Bam*HI (lane 1 and 6), *Eco*RV/*Bam*HI (lane 2), *Sma*I/*Bam*HI (lane 3), *Kpn*I/*Bam*HI (lane 4), *Cla*I/*Bam*HI (lane 5), hybridised to [α-$^{32}$P]dCTP-labelled pAN7-1. Lane M, 1 kb ladder.

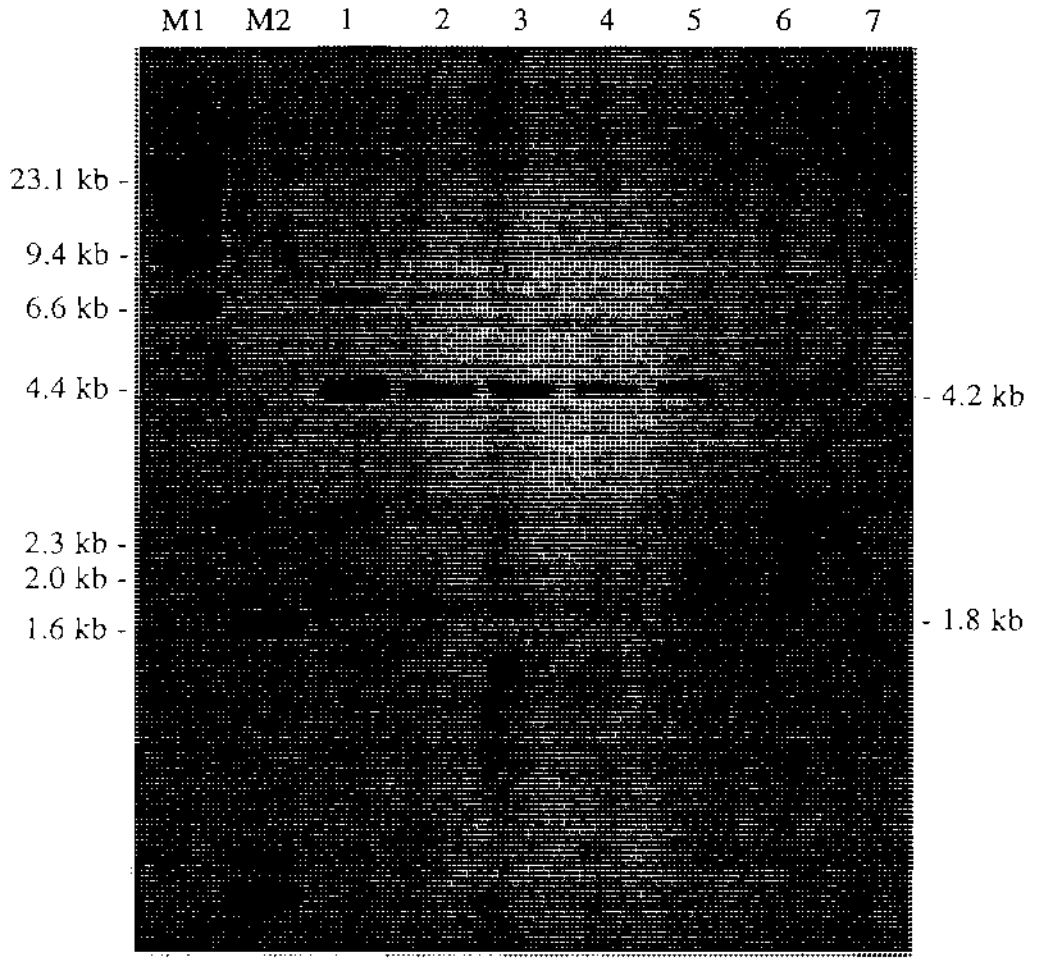



**Fig. 3.6** *Southern analysis of YI-20 DNA showing the pAN7-1 copy number.*

Southern analysis of *Eco*RI digested YI-20 DNA; 500 ng (lane 1), 250 ng (lane 2), 200 ng (lane 3), 150 ng (lane 4), 100 ng (lane 5), 50 ng (lane 6), 37.5 ng (lane 7), hybridised with the 0.8 kb *Bam*HI/*Eco*RI fragment (probe A) of pAN7-1. Lane M1, λ *Hin*dIII; lane M2, 1 kb ladder.

densitometry (with an Alpha Innotech Gel Documentation system) the intensity of both bands were determined and the area under the curve of the single copy band (1.8 kb) was directly compared to that of the multi-copy band (4.2 kb). The intensity of the 1.8 kb fragment was also compared to serial dilutions of the 4.2 kb multi-copy band to ensure that the linearity of the X-ray film had not been exceeded. Table 3.4 shows ratios of between 1:3 - 1:6, which indicates approximately 4 - 6 copies of pAN7-1 have integrated into YI-20. When comparing the area of the single-copy band in lane 1 to that of the $^1/_2$ and $^1/_5$ dilutions (lanes 2 and 5) a more consistent ratio of between 1:4.1 - 1:4.5 is obtained, indicating 5 - 6 copies are present.

## 3.3 *Isolation of Left-Hand Flanking Sequence From YI-20*

Analysis of the restriction enzyme map of YI-20 (Fig. 3.7) shows a 6.9 kb *Hin*dIII fragment containing the left-hand flanking sequence (pPN1375) could be plasmid rescued (Section 2.13). Such a fragment contains the origin of replication and an ampicillin-selectable marker required for maintenance in *E. coli*. However, the outcome of the plasmid rescue will also include intact pAN7-1 of the same size (6.8 kb). Self-ligated *Hin*dIII-digested YI-20 genomic DNA, was transformed into XL-1 cells (Section 2.13) and the resulting transformants screened by restriction enzyme analysis of DNA prepared by the rapid boil method (Section 2.5.3). Due to the similar size of the rescued plasmids (6.8 kb for pAN7-1 and 6.9 kb for pPN1375), the resulting clones were distinguished by digestion with five different restriction enzymes (Fig. 3.8). The *Hin*dIII digest allowed elimination of clones that had rescued non-related pieces of DNA. The *Bam*HI digest eliminated any pAN7-1 plasmids as pPN1375 does not contain a *Bam*HI site. The *Bgl*II, *Sau*3A and *Hae*III digests, cutting both pAN7-1 and pPN1375, give different banding patterns. Of the nine *E. coli* transformants isolated, one contained a partial copy of pAN7-1 which included the left-hand flanking sequence (pPN1375), while the remainder consisted of pAN7-1. A comparison of the restriction enzyme patterns between pAN7-1 and pPN1375 are shown in Fig. 3.8A and 3.8B. As expected, pPN1375 did not contain a *Bam*HI site (lane 4). However, restriction enzyme analysis using *Bgl*II showed an internal 0.5 kb *Bgl*II fragment (lane 6) that would be present in the wild-type genome. Previous data, seen in Fig. 3.4B, placed the first *Bgl*II site within the flanking sequence. Thus, the second site was unexpected and would only be detected with the isolation of flanking sequence. When digested with the 4 bp cutters, *Hae*III and *Sau*3A, a clear difference in banding patterns can be seen between the two plasmids. A shift from 0.75 kb (lane 1, pAN7-1) to 0.95 kb (lane 2, pPN1375) is seen in the *Hae*III digest, while the

*Table 3.4* *pAN7-1 copy number in YI-20*

| Lane | Band | Area of band | Ratio of band 2:1 |
|---|---|---|---|
| 1 | 1 | 2 048 | 1:3.2 |
| | 2 | 633 | |
| 2 | 1 | 1 447 | 1:4.3 |
| | 2 | 333 | |
| 3 | 1 | 1158 | 1:4.2 |
| | 2 | 275 | |
| 4 | 1 | 747 | 1:4.1 |
| | 2 | 184 | |
| 5 | 1 | 526 | 1:5.7 |
| | 2 | 92 | |
| 6 | 1 | 145 | |

[a]Refer to Fig. 3.6. Band 1 is the 4.2 kb multi-copy fragment and band 2 is the 1.8 kb single-copy fragment.

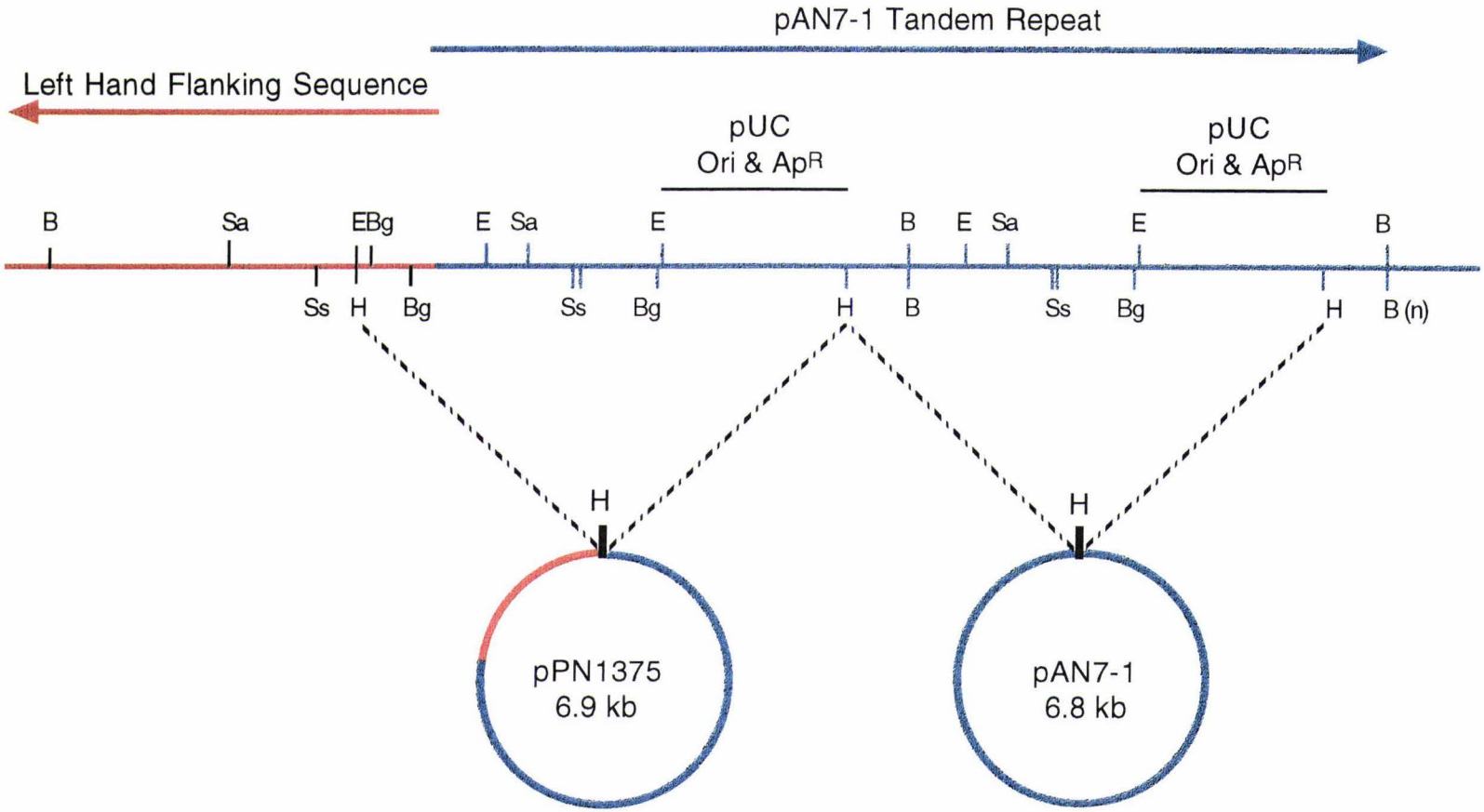



***Fig. 3.7*** *Restriction enzyme map of the pAN7-1 integration site in YI-20 showing the plasmid rescue outcomes with HindIII.*

Abbreviations are B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; Sa, *Sal*I; Ss, *Sst*I; pUC Ori & Ap$^R$, pUC18 sequence associated with the pAN7-1 vector that contains an *E. coli* origin of replication and an ampicillin-selectable marker.

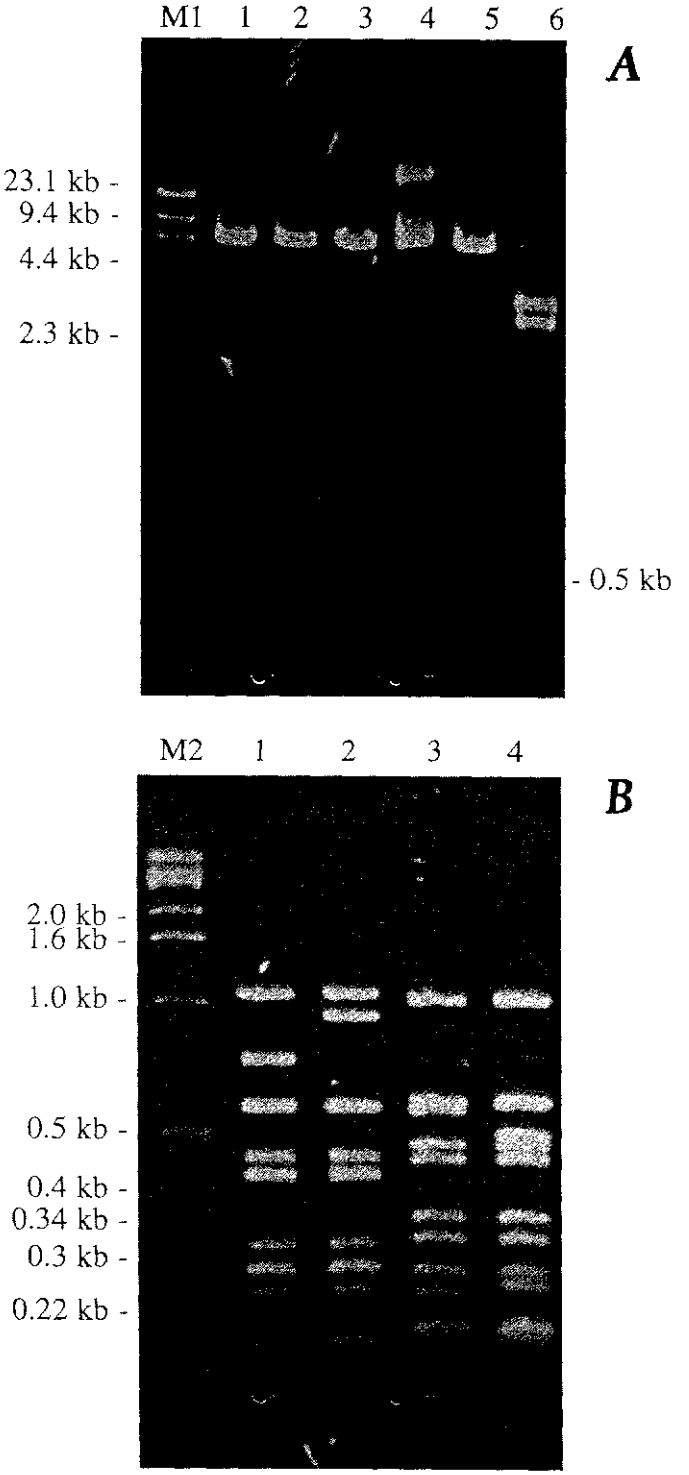



**Fig 3.8** *Restriction enzyme analysis to differentiate between pAN7-1 and pPN1375.*

***A*** *Hin*dIII digested pAN7-1 (lane 1), pPN1375 (lane 2). *Bam*HI digested pAN7-1 (lane 3), pPN1375 (lane 4). *Bgl*II digested pAN7-1 (lane 5), pPN1375 (lane 6). Lane M1, λ *Hin*dIII.

***B*** *Hae*III digested pAN7-1 (lane 1), pPN1375 (lane 2). *Sau*3A digested pAN7-1 (lane 3), pPN1375 (lane 4). Lane M2, 1 kb ladder.

*Sau3A* digest, shows a shift from 0.55 kb (lane 3, pAN7-1) to 0.49 kb (lane 4, pPN1375).

The internal 0.5 kb *Bgl*II fragment isolated from pPN1375 (Section 2.13) was [$\alpha$-$^{32}$P]dCTP-labelled and used to hybridise a Southern blot (Section 2.11) of wild-type genomic DNA (Fig. 3.9). The wild-type DNA showed the following hybridising bands: 1.6 kb *Eco*RI fragment (lane 2) also present in the *Hin*dIII/*Eco*RI digest (lane 1), 2.8 kb *Hin*dIII fragment (lane 3) and 0.5 kb *Bgl*II fragment (lane 4).

## 3.4 *Attempts to Isolate Right-Hand Flanking DNA*

To date I have been unable to establish the organisation of the right-hand flanking DNA. This is likely due to the difficulty of working with a tandem repeat, the head-to-head copy, and perhaps the presence of a partial copy of pAN7-1 adjacent to this side. Plasmid rescue (Section 2.13) from YI-20 was attempted in order to isolate right-hand flanking sequence using the following restriction enzymes, *Bam*HI, *Bgl*II, *Eco*RI, *Hin*dIII, *Sal*I and *Sst*I. These enzymes were chosen as they would potentially retain both the origin of replication and the ampicillin-selectable marker required for replication in *E. coli*. Self-ligation of YI-20 DNA digested with the above enzymes was performed as described in Section 2.13. Each of the resulting clones, (11 *Bam*HI, 7 *Bgl*II, 17 *Eco*RI, 6 *Hin*dIII, 1 *Sal*I and 4 *Sst*I) were first analysed with the original enzyme used in the plasmid rescue and then by two different 4 base pair cutters (*Hae*III and *Sau*3A). The *Hae*III and *Sau*3A digests of the clones isolated were checked on a 1.2 % (w/v) agarose gel (data not shown) and all of the ampicillin-resistant clones had the same banding pattern as that expected for pAN7-1 or its partial derivative.

## 3.5 *Molecular Analysis of the Wild-Type Region*

A $\lambda$-GEM11 (Appendix A1.4) library of wild-type (PN2013) *P. paxilli* (Itoh and Scott, 1994) was screened using the 0.5 kb *Bgl*II fragment from pPN1375 (Section 2.12). Two overlapping clones, $\lambda$CY1 and $\lambda$CY3, were isolated (Section 2.12). The average insert size was determined by *Sst*I digests where the cloned wild-type DNA is precisely removed from the arms. Insert sizes of 15.5 kb and 13.4 kb were found for $\lambda$CY1 and $\lambda$CY3 respectively (Fig 3.10, lane 3). Southern analysis of $\lambda$CY1 and $\lambda$CY3, using a range of restriction enzymes (Fig 3.10), determined that the DIG-labelled 0.5 kb *Bgl*II

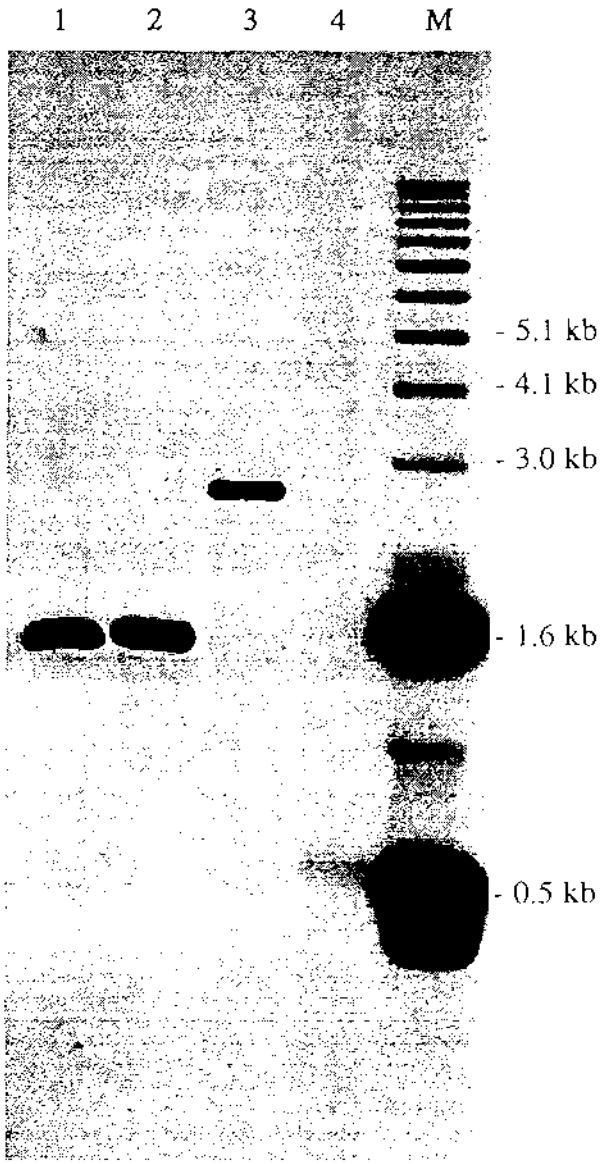



*Fig 3.9* *Southern analysis of wild-type DNA probed with resuced flanking sequence from pPN1375.*

Southern analysis of 1 µg of wild-type DNA digested with *Eco*RI/*Hin*dIII (lane 1), *Eco*RI (lane 2), *Hin*dIII (lane 3), *Bgl*II (lane 4), hybridised with the 0.5 kb *Bgl*II fragment from pPN1375. Lane M, 1 kb ladder.

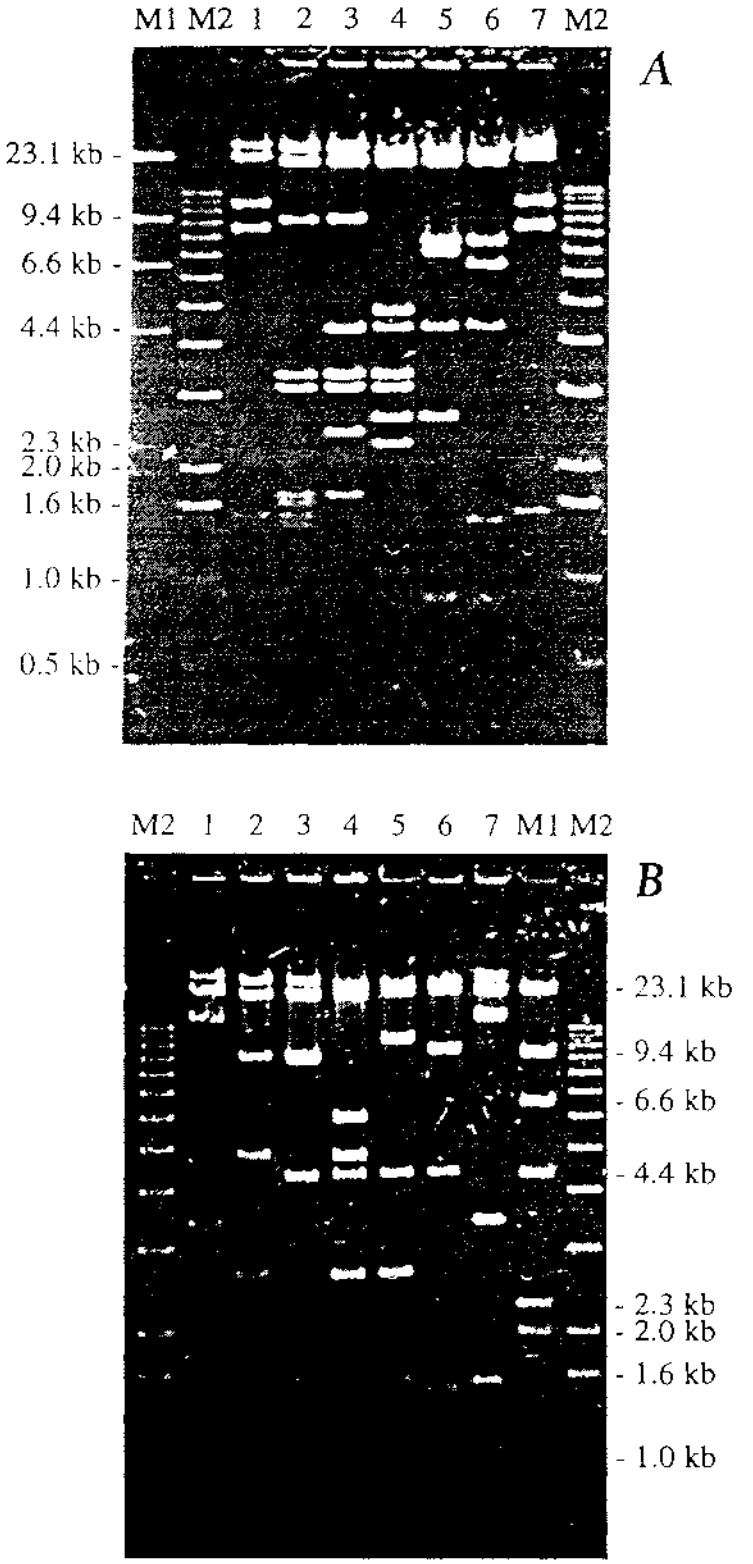



***Fig 3.10*** *Restriction enzyme digestion profiles of λCY1 and λCY3.*

λ clones digested with *Eco*RI (lanes 1 and 7), *Eco*RI/*Sst*I (lane 2), *Sst*I (lane 3), *Sst*I/*Hin*dIII (lane 4), *Hin*dIII (lane 5), *Hin*dIII/*Eco*RI (lane 6). Lane M1, λ *Hin*dIII. Lane M2, 1 kb ladder.

***A*** λCY1.

***B*** λCY3.

fragment (Section 2.11) hybridised to fragments of identical sizes in each λ clone (Fig. 3.11). The common hybridising bands are a 1.5 kb *Eco*RI fragment (lane 1 and 7), also seen in the *Eco*RI/*Sst*I digest (lane 2); a 4.5 kb *Sst*I fragment (lane 3); a 2.8 kb *Hin*dIII (lane 5), also seen in the *Sst*I/*Hin*dIII digest (lane 4); and a 1.4 kb *Hin*dIII/*Eco*RI fragment (lane 6). Other non-hybridising fragments seen in these two clones were different and are listed in Table 3.5.

### 3.5.1 *Sub-Cloning and Mapping of λCY-1 Fragments*

The following fragments; 1.5 kb *Eco*RI, 2.8 kb *Hin*dIII and 4.5 kb *Sst*I, were subcloned (Section 2.13) from λCY1 into pUC118 (pPN1385, pPN1387 and pPN1408, respectively). After purification of these plasmids by alkaline lysis and a CsCl density gradient (Section 2.5.5), these clones were used to devise a restriction enzyme map across the point of pAN7-1 integration (Fig. 3.12). In line with the hybridisation results (Fig. 3.11), digestion of the 2.8 kb *Hin*dIII clone (pPN1387) with *Eco*RI showed that it shared approximately 1.4 kb in common with the 1.5 kb *Eco*RI fragment (pPN1385) and a 0.6 kb fragment. As expected, the 4.5 kb *Sst*I fragment from pPN1408 contained the 1.5 kb and 0.6 kb *Eco*RI fragments, the 2.8 kb *Hin*dIII fragment, and a 0.7 kb *Bam*HI fragment. The 0.5 kb *Bgl*II fragment was contained within each of the three subclones (pPN1385, pPN1387 and pPN1408), while only pPN1387 and pPN1408 contained the 1.2 kb *Bgl*II fragment.

### 3.5.2 *Sequence Analysis of the 4.5 kb SstI Subclone (pPN1408)*

At this stage it was thought that sequence data obtained from these subcloned fragments would indicate the potential gene that was knocked out due to pAN7-1 integration and thus resulting in the Pax$^-$ phenotype. Sequence was originally obtained using the Sequenase kit (Section 2.14) with CsCl-purified DNA (Section 2.5.5) and the pUC forward and reverse primers. However, the above plasmids were also sequenced using the other methods stated in Section 2.14. To generate overlapping sequence data, primers (Section 2.15) were made to this region (Fig. 3.12) allowing extension of the sequence. The sequence was assembled using the GELASSEMBLE program of the GCG Wisconsin package (Section 2.15.5) and a consensus sequence generated (see Appendix A2.1 and A2.2 for BIGPIC and PRETTYOUT). Although this sequence has not been

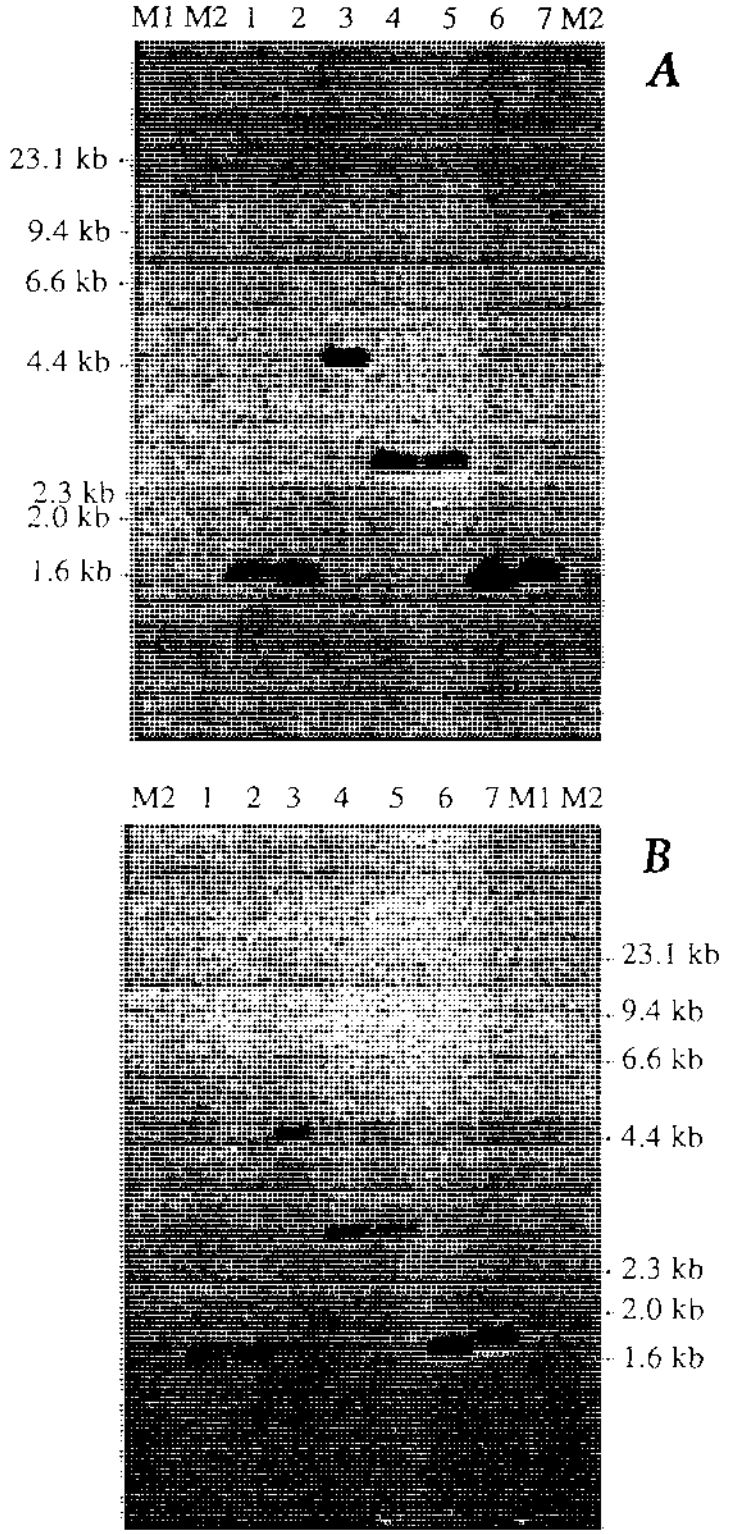



**Fig 3.11** *Mapping the position of the 0.5 kb BglII fragment on λCY1 and λCY3.*

Southern analysis of Fig. 3.10. λ clones digested with *Eco*RI (lanes 1 and 7), *Eco*RI/*Sst*I (lane 2), *Sst*I (lane 3), *Sst*I/*Hin*dIII (lane 4), *Hin*dIII (lane 5), *Hin*dIII/*Eco*RI (lane 6), hybridised with DIG labelled 0.5 kb *Bgl*II fragment from pPN1375. Lane M1, λ *Hin*dIII. Lane M2, 1 kb ladder.

***A*** λCY1.

***B*** λCY3.

*Table 3.5A* *Fragment sizes generated by single and double enzyme digestions of λCY1*

| *Eco*RI | *Eco*RI/*Sst*I | *Sst*I | *Sst*I/*Hin*dIII | *Hin*dIII | *Hin*dIII/*Eco*RI | *Eco*RI |
|---|---|---|---|---|---|---|
| 10.5 | 9.4 | 9.4 | 5.0 | 7.5 | 7.5 | 10.5 |
| 8.5 | 3.5 | **4.5** | 4.6 | 6.5 | 6.2 | 8.5 |
| **1.5** | 3.3 | 3.5 | 3.5 | 4.6 | 4.6 | **1.5** |
| | 1.7 | 3.3 | 3.3 | **2.8** | **1.4** | |
| | 1.6 | 2.5 | **2.8** | 0.9 | 0.9 | |
| | **1.5** | 1.7 | 2.3 | | 0.8 | |
| | 1.4 | | 1.1 | | 0.7 | |
| | | | 0.8 | | | |
| | | | 0.7 | | | |

*Table 3.5B* *Fragment sizes generated by single and double enzyme digestions of λCY3*

| *Eco*RI | *Eco*RI/*Sst*I | *Sst*I | *Sst*I/*Hin*dIII | *Hin*dIII | *Hin*dIII/*Eco*RI | *Eco*RI |
|---|---|---|---|---|---|---|
| 13.0 | 9.4 | 9.4 | 6.0 | 10.5 | 9.5 | 13.0 |
| 3.4 | 4.9 | 8.5 | 4.9 | 4.6 | 4.6 | 3.4 |
| **1.5** | 2.8 | **4.5** | 4.6 | **2.8** | **1.4** | **1.5** |
| 1.1 | **1.5** | 1.0 | **2.8** | 1.5 | 1.2 | 1.1 |
| | 1.4 | | 1.5 | 0.8 | 0.8 | |
| | | | 1.1 | 0.7 | 0.7 | |
| | | | 0.8 | | | |
| | | | 0.7 | | | |

Data from Fig. 3.10 and 3.11.

Fragments containing the left λ arm have been excluded from this table.

Fragment sizes in **bold** indicate hybridisation to the 0.5 kb *Bgl*II fragment from pPN1375.

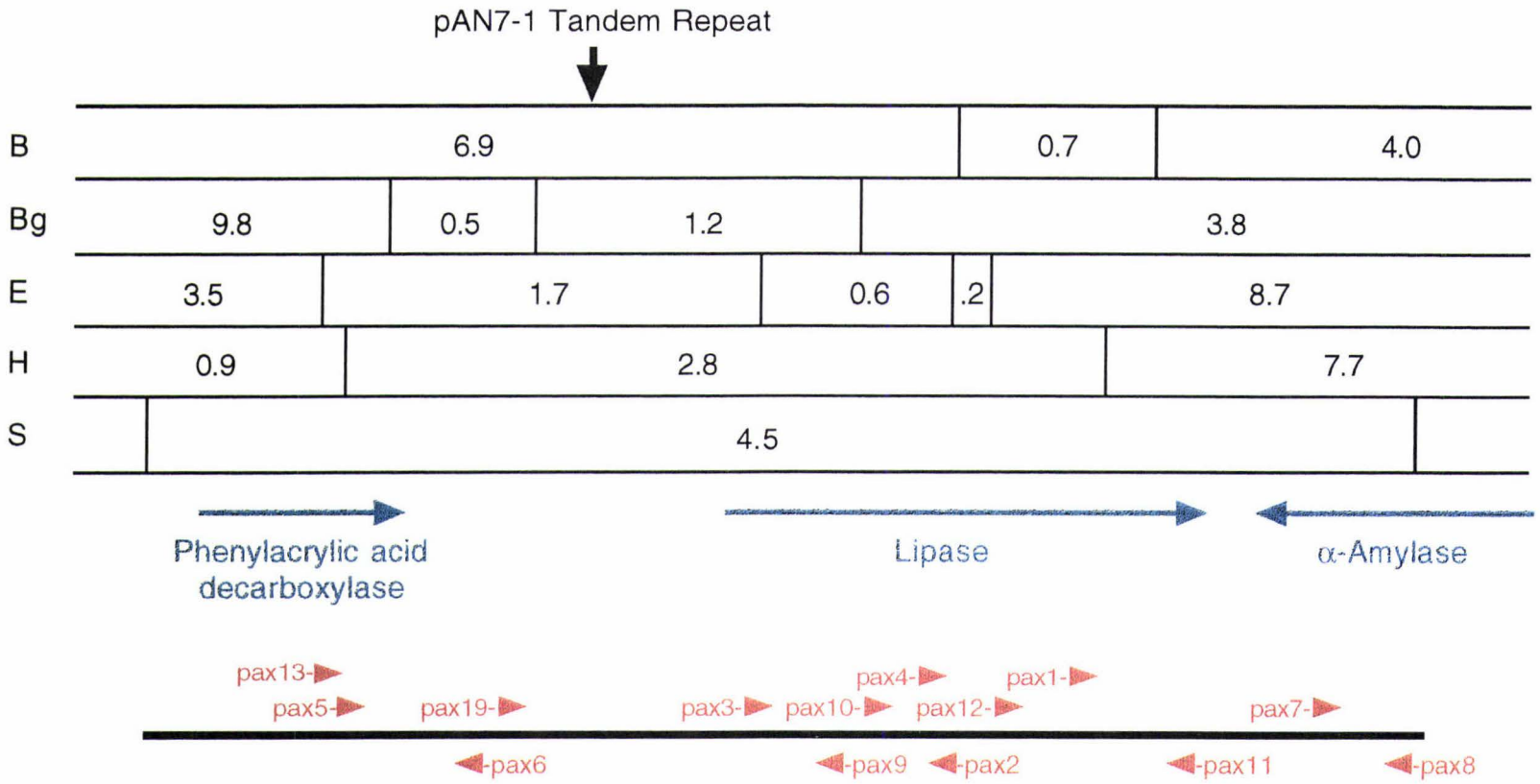



***Fig. 3.12*** *Restriction enzyme map of the 4.5 kb SstI fragment indicating gene and primer position.*

Abbreviations are B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; S, *Sst*I. Fragment sizes are shown in kilobases.

thoroughly edited, sequence analysis of the 4.5 kb *Sst*I fragment (pPN1408) using the FRAMES (Fig. 3.13) and MAP programs of the GCG Wisconsin package (Section 2.15.5) showed three distinctive open reading frames (ORFs). ORF 1 (Fig. 3.14) is seen in frame A (Fig. 3.13) and begins at nucleotide 199 with the stop codon at nucleotides 883 - 885. A BLAST search (Altschul *et al.*, 1990) of the protein databases (Brookhaven, SWISS-PROT and GenBank) with ORF 1 showed substantial similarity (a score of 598) to a phenylacrylic acid decarboxylase from *Saccharomyces cerevisiae* (accession number P33751). A BLAST search (Altschul *et al.*, 1990) of the protein databases with ORF 2, shows sequence similarity (a score of 855) to a triacylglycerol lipase from *Candida rugosa*, (accession number P32948). The correct start site of ORF 2 (Fig. 3.14), seen in frame C (Fig. 3.13), has not been determined as no potential start methionine is present at the expected location, this is likely due to a sequencing error across this region. However, the stop codon for ORF 2 is present at nucleotides 3642 - 3644. Due to the similarity of ORF 2 to the triacylglycerol lipase, the expected start codon would be nucleotides 1986 - 1988. Sequence analysis shows that ORF 3 (Fig. 3.14), seen in frame D (Fig. 3.13), is not a complete gene and must start in the adjacent fragment. A BLAST search (Altschul *et al.*, 1990) of this ORF against the protein databases, shows sequence similarity (a score of 263) to an acid-stable alpha-amylase from *Aspergillus kawachii* (accession number AB008370).

### 3.5.3 *Sequence Analysis Surrounding the pAN7-1 Point of Integration*

Mapping analysis places the pAN7-1 point of integration approximately 300 bp to the right of the 0.5 kb *Bgl*II fragment (Section 3.1.2). Sequence data of this region shows pAN7-1 has integrated at a site between genes (Fig. 3.14), and the surrounding genes seem unlikely to be the proteins responsible for the Pax$^-$ phenotype. To obtain sequence data across the pAN7-1 junction from pPN1375, pax19, a primer within the 0.5 kb BglII fragment of pPN1375, was made. Sequence data from this plasmid (Fig. 3.15) show pAN7-1 integration is between nucleotides 1506 and 1507 (Fig. 3.14 and 3.15A). This data confirm the loss of the *Bam*HI site at the point of pAN7-1 integration as eight bases, including this site, are missing at this junction (Fig. 3.15B). Previous work by Itoh and Scott (1994) had shown that deletions were a common outcome of pAN7-1 integration at a spore-pigment locus. Experiments were therefore carried out to check this possibility.

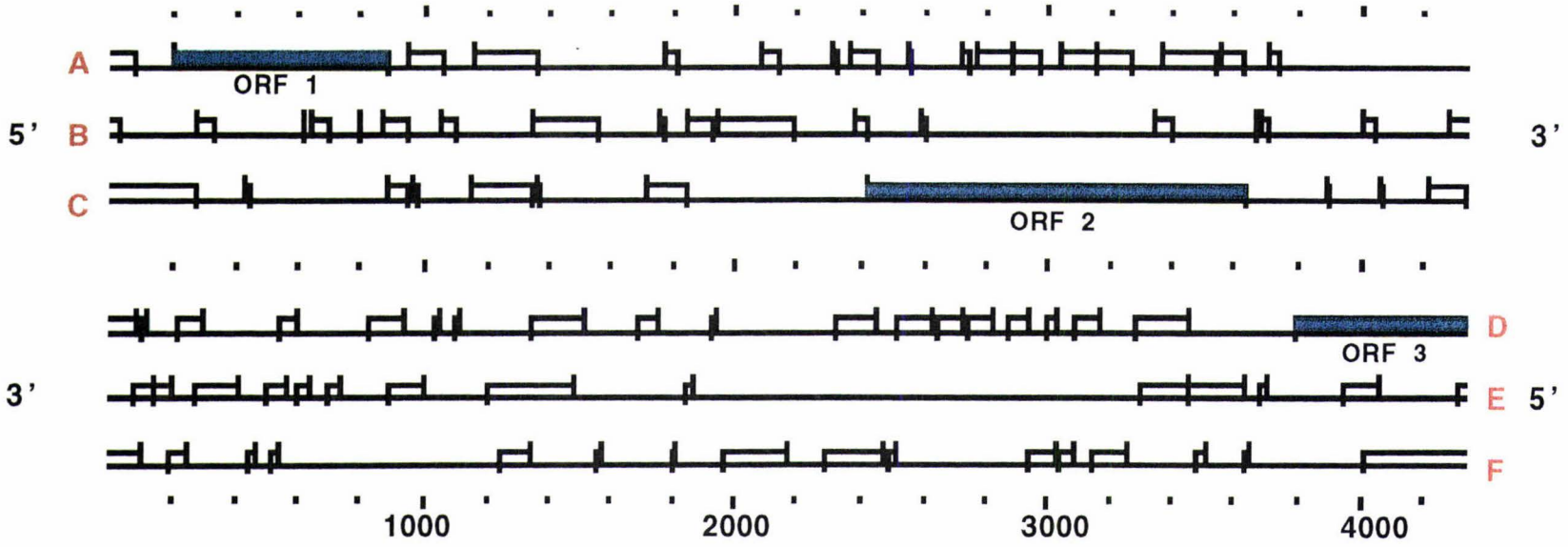


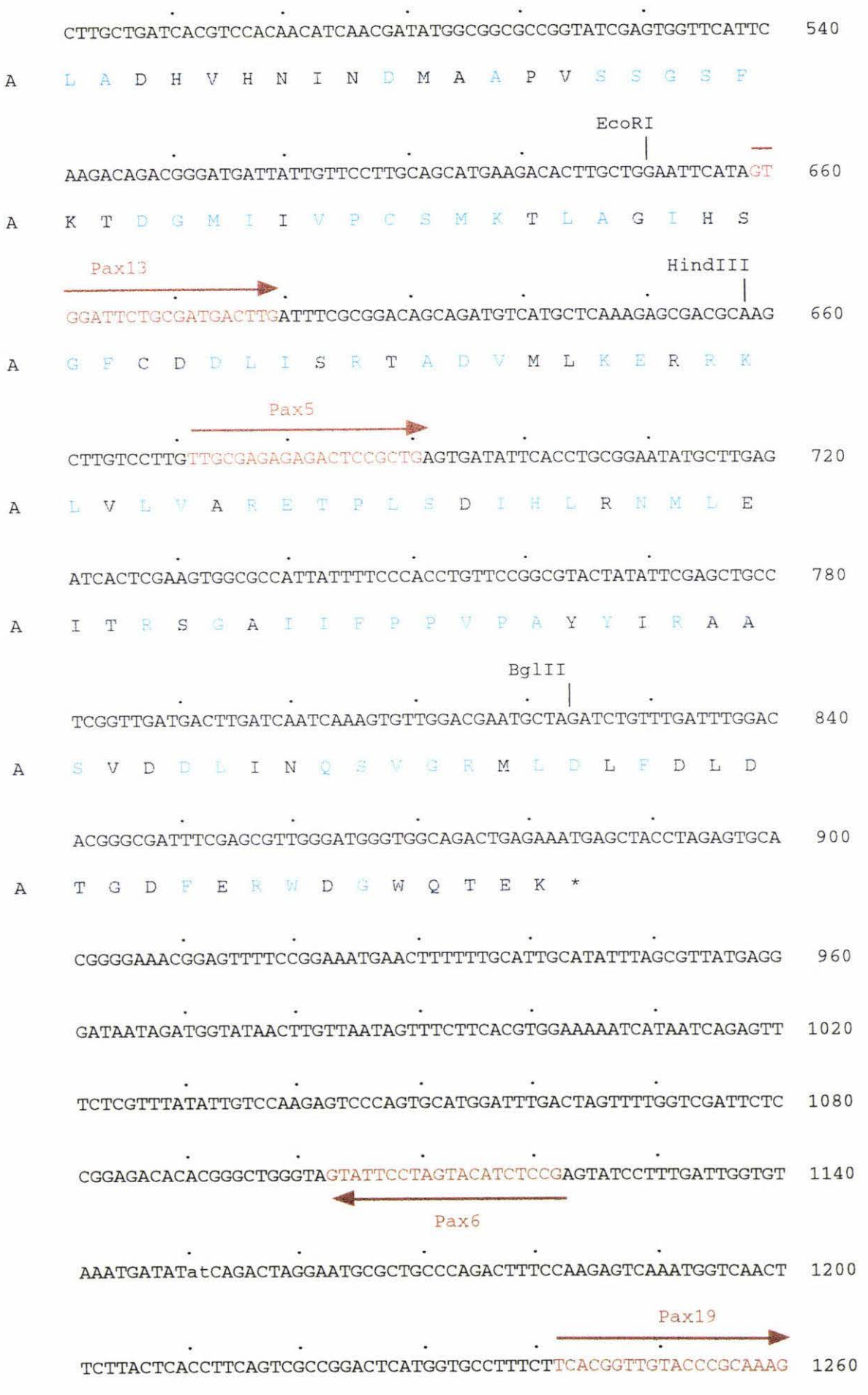

***Fig. 3.13*** *A 6-frame translation of the 4.5 kb SstI fragment from pPN1408.*

The FRAMES program of the GCG package plots the open reading frames (ORF) of a nucleic acid sequence as boxes bordered by potential start and stop codons. Potential start codons are shown as short lines that extend above the box and potential stop codons are shown as short lines that extend below the box. By default, only the start and stop codons at the ends of open reading frames are shown in the frame display; if a start codon is passed, no start codons are shown again until a stop codon is passed; if a stop codon has been passed, no stop codons are shown until a start codon is passed.

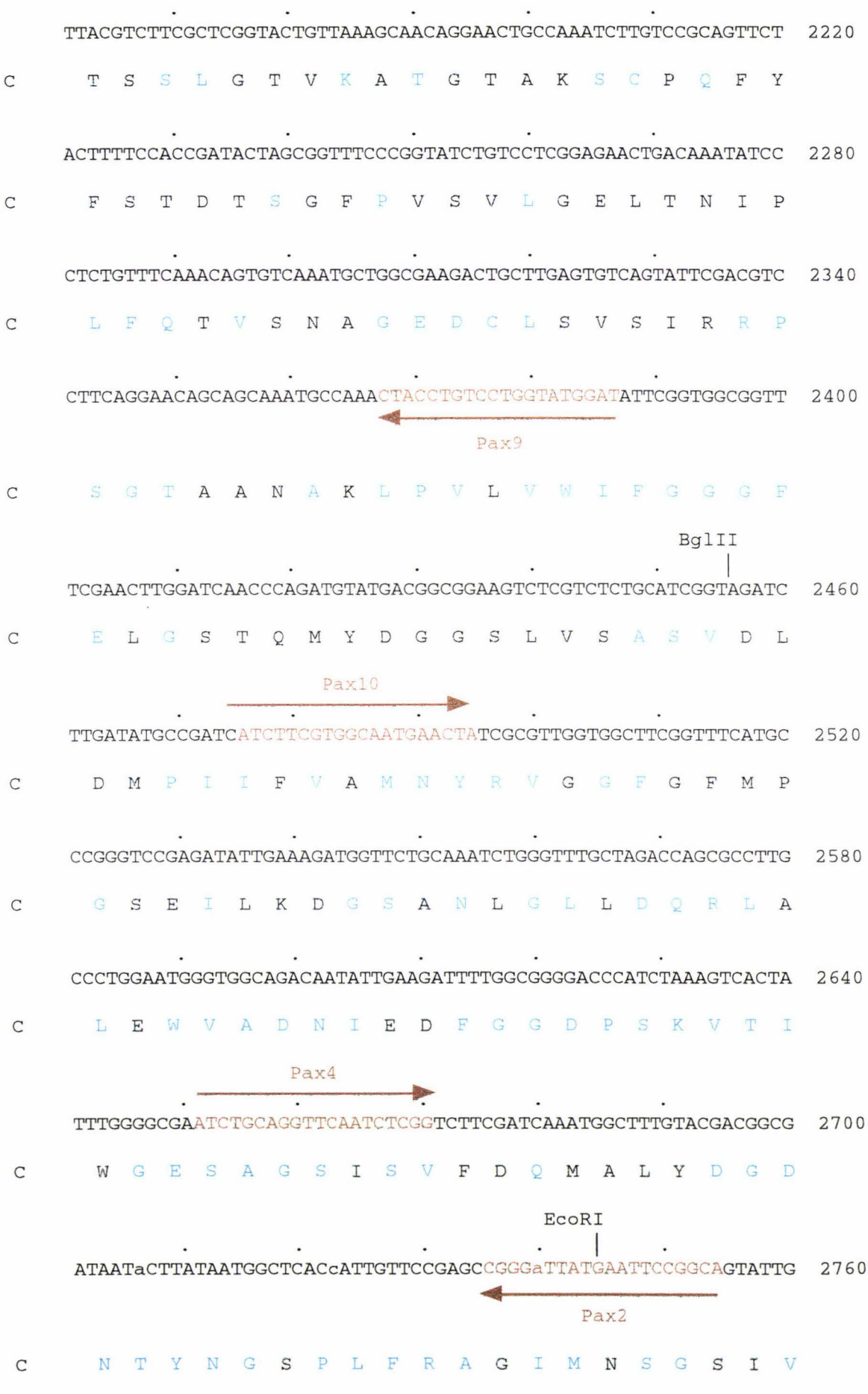
***Fig 3.14*** *DNA Sequence of the 4.5 kb SstI fragment and deduced amino acid sequence of the three open reading frames.*

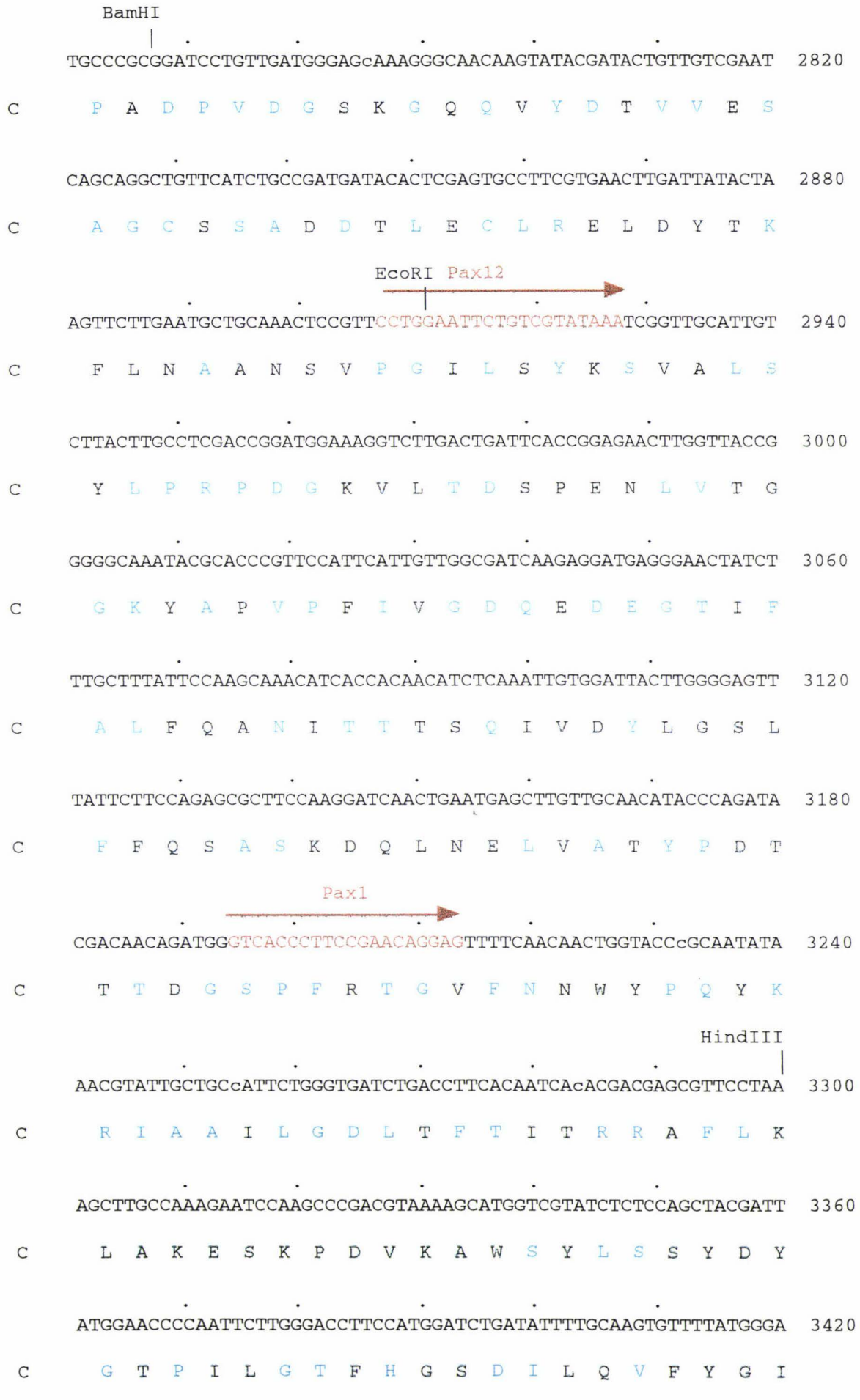
The 4,344 bp *Sst*I fragment from pPN1408 contains 3 open reading frames. ORF 1 codes from base 199 - 882, with sequence similarity to a phenylacrylic acid decarboxylase (Accession number P33751). The start of ORF 2 is not able to be positioned as no methionine is present in this area. However, similarity to a lipase (Accession number P32948) indicates it begins near base 1986 with the stop codon at bases 3,642 - 3,644. ORFs 1 and 2 are translated in the forward frames. ORF 3 is not complete begining in the adjacent *Sst*I fragment (pPN1422). ORF 3 is translated in the reverse frame with sequence similarity to an acid-stable alpha-amylase (Accession number AB008370). No introns have been detected in these three ORFs. Sequence identity of each ORF to their respective genes are shaded light blue, the remaining amino acids are shaded dark blue. Primers are indicated in red.

```
 SstI
    |    .         .         .         .         .
GAGCTCGCCGAGTTTCTGTTTGCCACGAATGGATATATAACATCCAGTCTTTGCTTGAAG     60

         .         .         .         .         .
TTGCAGTAATCTTGCAATCATCGATCAAAAGCATCACTGCACCTTGTCTTGCATCCACCA    120

         .         .         .         .         .
CTGGCAGTACTCCACATCTATCACACTACTACTGGATCAGCAACTCTACAAACCTGAGAA    180

         .         .         .         .         .
AGACGACTTACCTCCGCCATGCTGTCTTTACTCTCAGGAGCACAAACGCCCCAATCCGGA    240

A                    M  L  S  L  L  S  G  A  Q  T  P  Q  S  G

         .         .         .         .         .
CGCACATCGCCTGGCAGTCACCACGATCATGAAACCTCCCACAACGAAACCCGCCATGTC    300

A    R  T  S  P  G  S  H  H  D  H  E  T  S  H  N  E  T  R  H  V

         .         .         .         .         .
TCAAGTCCACCCCGCAAGCGCATTGTGGTTGCCATGACCGGTGCCACCGGCGCAATGCTT    360

A    S  S  P  P  R  K  R  I  V  V  A  M  T  G  A  T  G  A  M  L

         .         .         .         .         .
GGTATCAAAGTCCTTATCGCCCTCCGCCGCCTCAACATCGAAACCCACCTGATCATGAGC    420

A    G  I  K  V  L  I  A  L  R  R  L  N  I  E  T  H  L  I  M  S

         .         .         .         .         .
AAATGGGCCGAAGCAACAATTAAATACGAGACCGATTATCACCCATCAAACGTCAAGGCC    480

A    K  W  A  E  A  T  I  K  Y  E  T  D  Y  H  P  S  N  V  K  A
```


CTTGCTGATCACGTCCACAACATCAACGATATGGCGGCGCCGGTATCGAGTGGTTCATTC 540
A L A D H V H N I N D M A A P V S S G S F
EcoRI
AAGACAGACGGGATGATTATTGTTCCTTGCAGCATGAAGACACTTGCTGGAATTCATAGT 660
A K T D G M I I V P C S M K T L A G I H S
Pax13
HindIII
GGATTCTGCGATGACTTGATTTCGCGGACAGCAGATGTCATGCTCAAAGAGCGACGCAAG 660
A G F C D D L I S R T A D V M L K E R R K
Pax5
CTTGTCCTTGTTGCGAGAGAGACTCCGCTGAGTGATATTCACCTGCGGAATATGCTTGAG 720
A L V L V A R E T P L S D I H L R N M L E
ATCACTCGAAGTGGCGCCATTATTTTCCCACCTGTTCCGGCGTACTATATTCGAGCTGCC 780
A I T R S G A I I F P P V P A Y Y I R A A
BglII
TCGGTTGATGACTTGATCAATCAAAGTGTTGGACGAATGCTAGATCTGTTTGATTTGGAC 840
A S V D D L I N Q S V G R M L D L F D L D
ACGGGCGATTTCGAGCGTTGGGATGGGTGGCAGACTGAGAAATGAGCTACCTAGAGTGCA 900
A T G D F E R W D G W Q T E K *
CGGGGAAACGGAGTTTTCCGGAAATGAACTTTTTTGCATTGCATATTTAGCGTTATGAGG 960
GATAATAGATGGTATAACTTGTTAATAGTTTCTTCACGTGGAAAAATCATAATCAGAGTT 1020
TCTCGTTTATATTGTCCAAGAGTCCCAGTGCATGGATTTGACTAGTTTTGGTCGATTCTC 1080
CGGAGACACACGGGCTGGGTAGTATTCCTAGTACATCTCCGAGTATCCTTTGATTGGTGT 1140
Pax6
AAATGATATatCAGACTAGGAATGCGCTGCCCAGACTTTCCAAGAGTCAAATGGTCAACT 1200
Pax19
TCTTACTCACCTTCAGTCGCCGGACTCATGGTGCCTTTCTTCACGGTTGTACCCGCAAAG 1260


CTACTTGTCTGCAGTCTGTCAGGAACAAGGGGCTTGCCAATAAAATGGGCTTCGTGTGGA 1320
BglII
AGATCAAGATCTTGCCACTCATTCACCGCAAATGAATATCCTGGATGAGATGATGGTGTG 1380
TAGCAACAAGCGAACCGTGCCCACTGGGGGATCTGCAGCTCTACCCTTCTGGGAAGCGGC 1440
ATCTTCCCCACCCCCAAGATTCGCGGGGCTGATGCGGTCATCGTGCATTCCACAGTTCGT 1500
pAN7-1 Insertion point
GCCCTCCGGATATCTGGTCATCCGACACCACTGGGTGCCGATTTCTACCCTATaAGCAGT 1560
CTGACATTGGAACCCGCAAAACGGcgTTACGCCTTGGCCTGCTCGTTCGCCAACCAGGAT 1620
CGAACTTGTCAAGTCCTTCACACTTCGCGAGTCTTGATATACCCAATAATAACTTTCTCA 1680
CGAACTCTCCGGCGAACTTTCCGTTCTTAGTATAGATGCTACTTCAACTGACTGTCGAAG 1740
ACCCCTCTTCTTCATGGCGGGGTCGGAATGAACCTTGCGATGTCTCTCTTCAACCGAGGA 1800
ACCACATGTCTCTGATCTTGATACGGGATTTTTCAAGAATCCGATGATGTGGAGAGGAGG 1860
TTATAATCCATCCCTTGGTCCCTTGAGGGTTGAAAAAGGGCGCTCCAGGACCCAAACGAG 1920
ACTCGGCTAGATCAAGATCAACATGCTGTTCGGGCGgATCAGTCCTCTACTTCCTGTCTT 1980
CTCTCACACTGGTGTTTGCTGCACCGGCGGGACAAAAGAGAGcAAGTGCCCCTACAGTCA 2040
C T L V F A A P A G Q K R A S A P T V T
Pax3
CCATATCGAAGCCTGCTGCTACGGTTATTGGATCTTCTTCAAACAATGTGGACAGCTTcA 2100
C I S K P A A T V I G S S S N N V D S F N
EcoRI
ATGGAATTCCCTTTGCGCAGGCACCGACCGGGTCACTCCGACTAAAGCCTCCAAAGGCCA 2160
C G I P F A Q A P T G S L R L K P P K A I

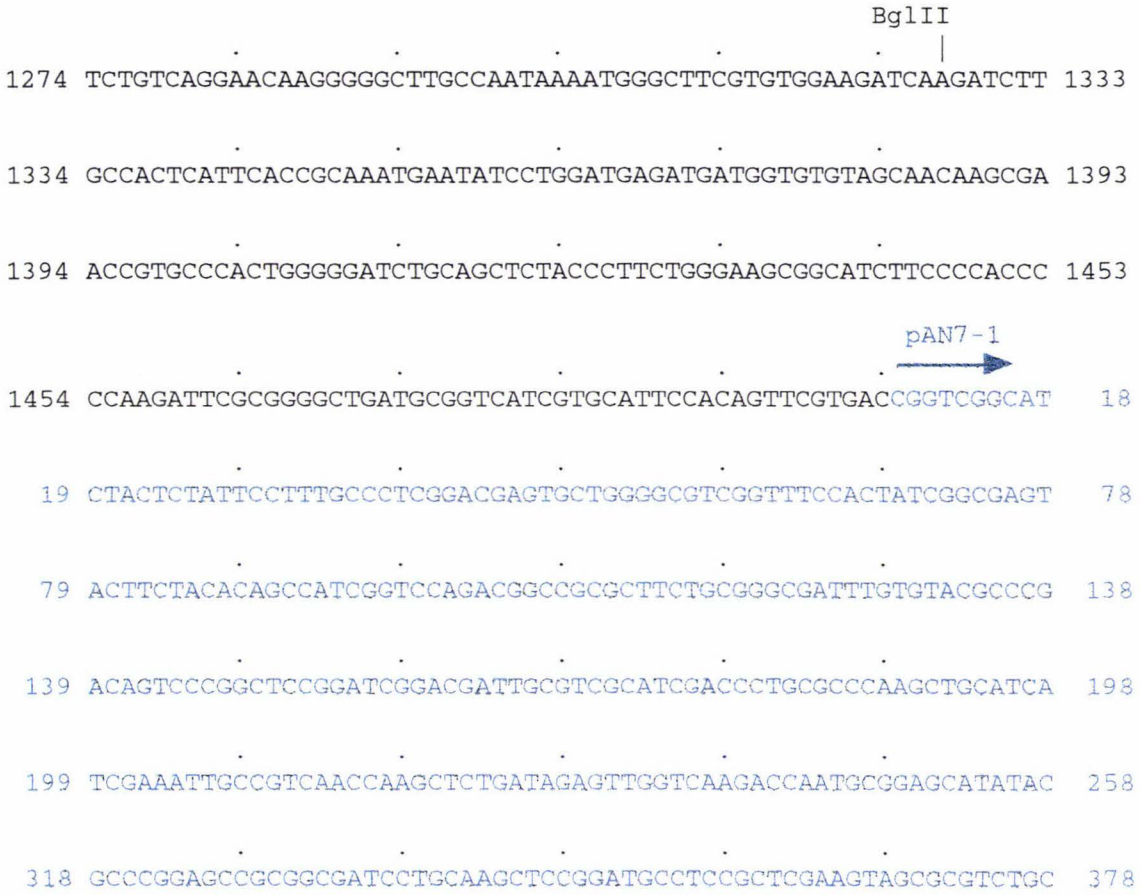

TTACGTCTTCGCTCGGTACTGTTAAAGCAACAGGAACTGCCAAATCTTGTCCGCAGTTCT 2220
C T S S L G T V K A T G T A K S C P Q F Y
ACTTTTCCACCGATACTAGCGGTTTCCCGGTATCTGTCCTCGGAGAACTGACAAATATCC 2280
C F S T D T S G F P V S V L G E L T N I P
CTCTGTTTCAAACAGTGTCAAATGCTGGCGAAGACTGCTTGAGTGTCAGTATTCGACGTC 2340
C L F Q T V S N A G E D C L S V S I R R P
CTTCAGGAACAGCAGCAAATGCCAAACTACCTGTCCTGGTATGGATATTCGGTGGCGGTT 2400
Pax9
C S G T A A N A K L P V L V W I F G G G F
BglII
TCGAACTTGGATCAACCCAGATGTATGACGGCGGAAGTCTCGTCTCTGCATCGGTAGATC 2460
C E L G S T Q M Y D G G S L V S A S V D L
Pax10
TTGATATGCCGATCATCTTCGTGGCAATGAACTATCGCGTTGGTGGCTTCGGTTTCATGC 2520
C D M P I I F V A M N Y R V G G F G F M P
CCGGGTCCGAGATATTGAAAGATGGTTCTGCAAATCTGGGTTTGCTAGACCAGCGCCTTG 2580
C G S E I L K D G S A N L G L L D Q R L A
CCCTGGAATGGGTGGCAGACAATATTGAAGATTTTGGCGGGGACCCATCTAAAGTCACTA 2640
C L E W V A D N I E D F G G D P S K V T I
Pax4
TTTGGGGCGAATCTGCAGGTTCAATCTCGGTCTTCGATCAAATGGCTTTGTACGACGGCG 2700
C W G E S A G S I S V F D Q M A L Y D G D
EcoRI
ATAATaCTTATAATGGCTCACcATTGTTCCGAGCCGGGaTTATGAATTCCGGCAGTATTG 2760
Pax2
C N T Y N G S P L F R A G I M N S G S I V


BamHI
TGCCCGCGGATCCTGTTGATGGGAGcAAAGGGCAACAAGTATACGATACTGTTGTCGAAT 2820
C P A D P V D G S K G Q Q V Y D T V V E S
CAGCAGGCTGTTCATCTGCCGATGATACACTCGAGTGCCTTCGTGAACTTGATTATACTA 2880
C A G C S S A D D T L E C L R E L D Y T K
EcoRI Pax12
AGTTCTTGAATGCTGCAAACTCCGTTCCTGGAATTCTGTCGTATAAATCGGTTGCATTGT 2940
C F L N A A N S V P G I L S Y K S V A L S
CTTACTTGCCTCGACCGGATGGAAAGGTCTTGACTGATTCACCGGAGAACTTGGTTACCG 3000
C Y L P R P D G K V L T D S P E N L V T G
GGGGCAAATACGCACCCGTTCCATTCATTGTTGGCGATCAAGAGGATGAGGGAACTATCT 3060
C G K Y A P V P F I V G D Q E D E G T I F
TTGCTTTATTCCAAGCAAACATCACCACAACATCTCAAATTGTGGATTACTTGGGGAGTT 3120
C A L F Q A N I T T T S Q I V D Y L G S L
TATTCTTCCAGAGCGCTTCCAAGGATCAACTGAATGAGCTTGTTGCAACATACCCAGATA 3180
C F F Q S A S K D Q L N E L V A T Y P D T
Pax1
CGACAACAGATGGGTCACCCTTCCGAACAGGAGTTTTCAACAACTGGTACCcGCAATATA 3240
C T T D G S P F R T G V F N N W Y P Q Y K
HindIII
AACGTATTGCTGCcATTCTGGGTGATCTGACCTTCACAATCAcACGACGAGCGTTCCTAA 3300
C R I A A I L G D L T F T I T R R A F L K
AGCTTGCCAAAGAATCCAAGCCCGACGTAAAAGCATGGTCGTATCTCTCCAGCTACGATT 3360
C L A K E S K P D V K A W S Y L S S Y D Y
ATGGAACCCCAATTCTTGGGACCTTCCATGGATCTGATATTTTGCAAGTGTTTTATGGGA 3420
C G T P I L G T F H G S D I L Q V F Y G I


TTTTGCCAAACTACGCGTCCAAGTCTTTCCATTCATACTATCTTTCTTTTGTCTATGGTA 3480
C L P N Y A S K S F H S Y Y L S F V Y G M
BamHI
TGGATCCCAATGCAAGGGCAACGGACTTCATGGATTGGCCAGAGTGGGGATCGAATCAAA 3540
C D P N A R A T D F M D W P E W G S N Q T
CACTGATGCAGTTCTTCAATGATCGAGGTGCACTTCTTGCGGATAACTTTCGACAGGATA 3600
Pax11
C L M Q F F N D R G A L L A D N F R Q D T
CTTATGATTTTATTCTGGAGAATGTAGGGTCTTTCCATATCTAGGAAGCCAACATTTAAT 3660
C Y D F I L E N V G S F H I *
AATGAGATATTGATCAATGTTTACAATGATGAATAATTGCCCACGAAATTCATGAATTTG 3720
GAGTTGAAGTAGTCAAACAAAGGTACTAATCTTCAGTATCTGAAGCTGACTTGAACTAGT 3780
GATATCTAGCAGCTATCGCCAGGTCGCAGTAACGGTAGCTGTAGATCCGGAGCAGCCAGT 3840
D * R W T A T V T A T S G S C G T
GGGGACGGTGTAGGACCTGTTCGGATCGCTCTCCCAAGTCACCGATGAAGACCCACTCGA 3900
D P V T Y S R N P D S E W T V S S S G S S
CTTTTCAATAAACTTATACTCAAAAGACGAGCCAACGGGGATAGTCACCTCAACATACCA 3960
D K E I F K Y E F S S G V P I T V E V Y W
Pax7
AAGTGGATTGGAAGAGGTATACTGATCTGCAGAAAGAGCAATGGCATCATCGGTGTCCCA 4020
D L P N S S T Y Q D A S L A I A D D T D W
GCTACCCAACTGGCTGATTGAGCCTGAGAGGTAAACATTTTGGCCATATGAGGTGGTGAC 4080
D S G L Q S I S G S L Y V N Q G Y S T T V


GACCTCTTTGAACAGCACGGGTAAGGCAGTGGCCTGGGTACAGGTTGAAGTGGTACTGGC 4140
D V E K F L V P L A T A Q T C T S T T S A
ACCGCCGGTACTTGAGGTCTTTGAACTGGTGCTCGTGGTTGACGCGCTTGTCGTAGTTCC 4200
D G G T S S T K S S T S T T S A S T T T G
GGAGTCTGAAGATGCACAGGGTCCACTGCTCTTCGCGGCAGAAGCAAGCATCAACACTCG 4260
D S D S S A C P G S S K A A S A L M L V R
AGGCAAGCCAGAGACCATTGGGACAGCGATGTTGCCGCTCGAGTCGACTGTTATGGATGT 4320
Pax8
D P L G S V M P V A I N G S S D V T I S T
SstI
ACATGTGTACATTTCCATGAGCTC 4344
D C T Y M E M L E

*A*




*B*


***Fig 3.15*** *DNA Sequence across the pAN7-1 point of integration in pPN1375.*

***A*** Using the primer pax 19 with pPN1375, the DNA sequence was obtained across the point of pAN7-1 integration. The black sequence is wild-type and is numbered as seen in Fig. 3.14. The blue sequence is pAN7-1 and is numbered from the *Bam*HI site.

***B*** Sequence of pAN7-1 starting at the *Bam*HI site. The blue pAN7-1 sequence is present in pPN1375. The red sequence is missing from the point of pAN7-1 integration.

Hybridisation of [α-$^{32}$P]dCTP-labelled λCY1 (Section 2.11) to *Bam*HI, *Bgl*II, *Eco*RI, *Hin*dIII and *Sst*I digests of PN2013 and YI-20 DNA showed that a number of hybridising bands present in wild-type are missing in the mutant (Fig. 3.16). Only fragments that map to the left of the pAN7-1 integration site in YI-20 hybridised to the λCY1 probe. The hybridising bands from Fig. 3.16 are described in Table 3.6. The hybridising bands common to both wild-type and YI-20 are the 9.8 and 0.5 kb *Bgl*II fragments (lanes 2 and 7), the 3.5 kb *Eco*RI fragment (lanes 3 and 8), the 0.9 and 0.8 kb *Hin*dIII fragments (lanes 4 and 9), and the 9.4 kb *Sst*I fragment (lanes 5 and 10). The following fragments are absent from YI-20: 6.9, 5.5, 4.0, 3.1, 1.2, 0.7 and 0.3 kb *Bam*HI (lanes 1 and 6); 4.6, 3.8, 2.1 and 1.2 kb *Bgl*II (lanes 2 and 7); 8.7, 5.2, 1.7, 0.6 and 0.2 kb *Eco*RI (lanes 3 and 8); 7.7, 2.8 and 2.5 kb *Hin*dIII (lanes 4 and 9); and 10.6, 4.5, 3.5 and 3.3 kb *Sst*I (lanes 5 and 10). YI-20 fragments that correspond to the left-hand junction fragments of the pAN7-1 insertion are as follows: 12 kb *Bam*HI (lane 6), 6.7 kb *Bgl*II (lane 7), 2.7 kb *Eco*RI (lane 8), 6.9 kb *Hin*dIII (lane 9) and 4.2 kb *Sst*I (lane 10). These results show that YI-20 has at least a 12 kb deletion from the point of pAN7-1 integration.

Data from this Southern result showed that most of λCY1 is contained within the deleted region, while most of λCY3 maps to the left of this deletion. These clones contain identical hybridising fragments as those found in genomic digests of wild-type indicating that there were no cloning artefacts. This result and subcloning of the *Sst*I fragments enabled a map (Fig. 3.17) to be drawn using the 5 enzymes used in the Southern analysis.

The following fragments were subcloned into pUC118 (Appendix A1.2): 2.5 kb (a truncated 10.6 kb wild-type fragment), 3.3 kb and 3.5 kb *Sst*I fragments (pPN1406, pPN1421 and pPN1422, respectively). The fragments were spot sequenced (Section 2.14) using the pUC forward and reverse primers (Section 2.15). The sequence data generated (see Appendix A2.3, A2.4 and A2.5) were used for BLAST searches (Altschul *et al.*, 1990) against the protein databases, Brookhaven, SWISS-PROT and GenBank. The 4.5 kb and 3.5 kb *Sst*I fragments spanning an *Sst*I junction (Fig. 3.17) showed sequence similarities to an acid stable alpha-amylase gene from *Aspergillus kawachii* (accession number AB008370). The 3.5 kb and 3.3 kb *Sst*I fragments spanning a second *Sst*I junction (Fig. 3.17) had sequence similarity to a glucoamylase gene from *Aspergillus awamori* (accession number U59303). The 3.3 kb and 2.5 kb (a truncated 10.6 kb wild-type fragment) *Sst*I fragments spanning a third *Sst*I junction showed

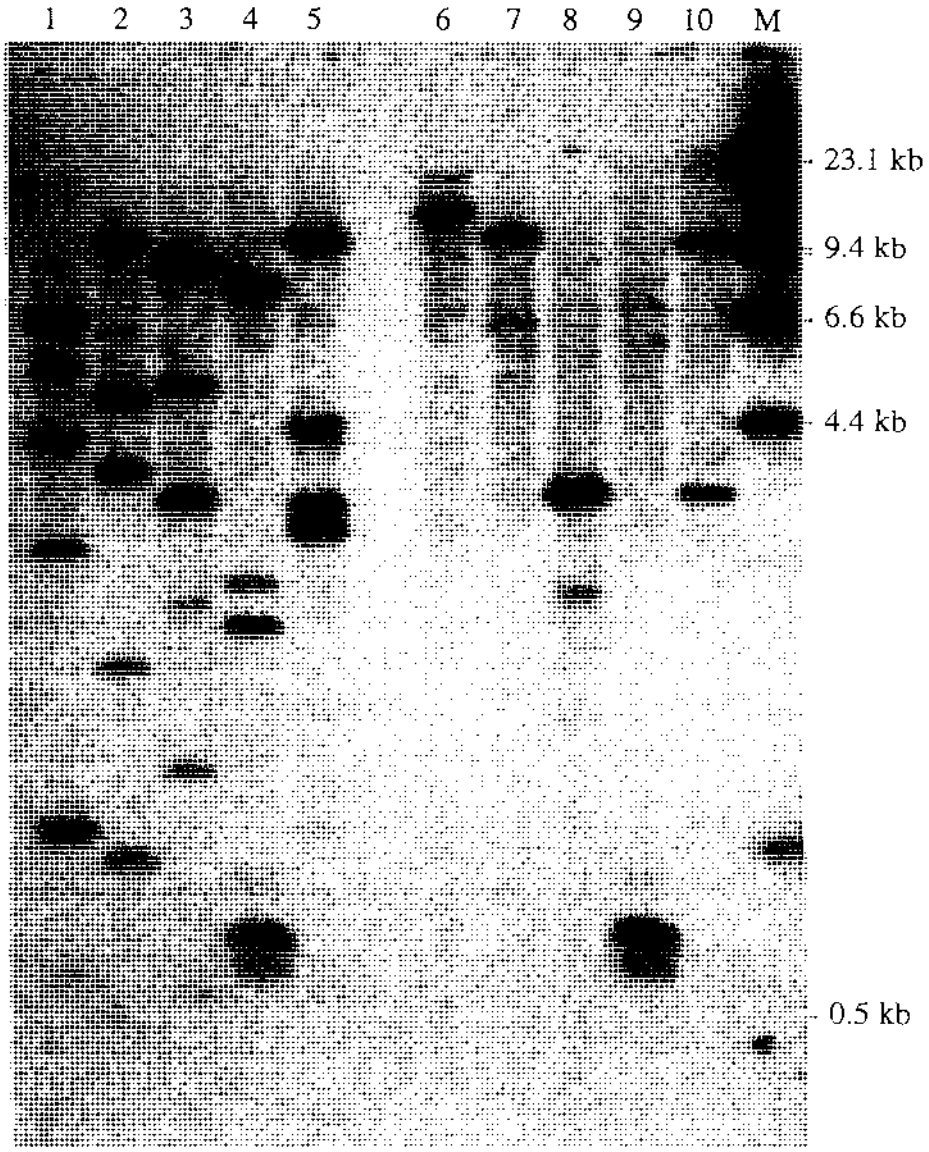



**Fig. 3.16** *Hybridisation of λCY1 to genomic DNA digests of wild-type and YI-20.*

Southern analysis of 1 μg of wild-type DNA (lanes 1-5) and YI-20 DNA (lanes 6-10) digested with *Bam*HI (lanes 1 and 6), *Bgl*II (lanes 2 and 7), *Eco*RI (lanes 3 and 8), *Hin*dIII (lanes 4 and 9), *Sst*I (lanes 5 and 10), hybridised with [α-$^{32}$P]dCTP-labelled λCY1. Lane M, λ*Hin*dIII.

**Table 3.6** *Hybridising bands common to wild-type and YI-20.*

| Restriction Enzyme | Wild-type kb | YI-20 kb |
|---|---|---|
| *Bam*HI | 6.9 | 12 |
| | 0.7 | |
| | 4.0 | |
| | 0.3* | |
| | 1.2 | |
| | 3.1 | |
| | 5.5 | |
| *Bgl*II | 9.8 | 9.8 |
| | 0.5 | 0.5* |
| | 1.2 | 6.7 |
| | 2.1 | |
| | 3.8 | |
| | 4.6 | |
| | 9.5* | |
| *Eco*RI | 3.5 | 3.5 |
| | 1.7 | 2.7 |
| | 0.6 | |
| | 0.2* | |
| | 8.7 | |
| | 5.2 | |
| *Hin*dIII | 0.8 | 0.8 |
| | 0.9 | 0.9 |
| | 2.8 | 6.8 |
| | 7.7 | |
| | 2.5 | |
| *Sst*I | 9.4 | 9.4 |
| | 4.5 | 4.2 |
| | 3.5 | |
| | 3.3 | |
| | 10 | |

Hybridising fragments common to wild-type and YI-20 are in red.
Hybridising fragments specific to YI-20 due to pAN7-1 integration are in blue.
Hydridising fragments that are only present in the wild-type are in black.

* These fragments are not visible in Fig. 3.16 but have been confirmed by other means.

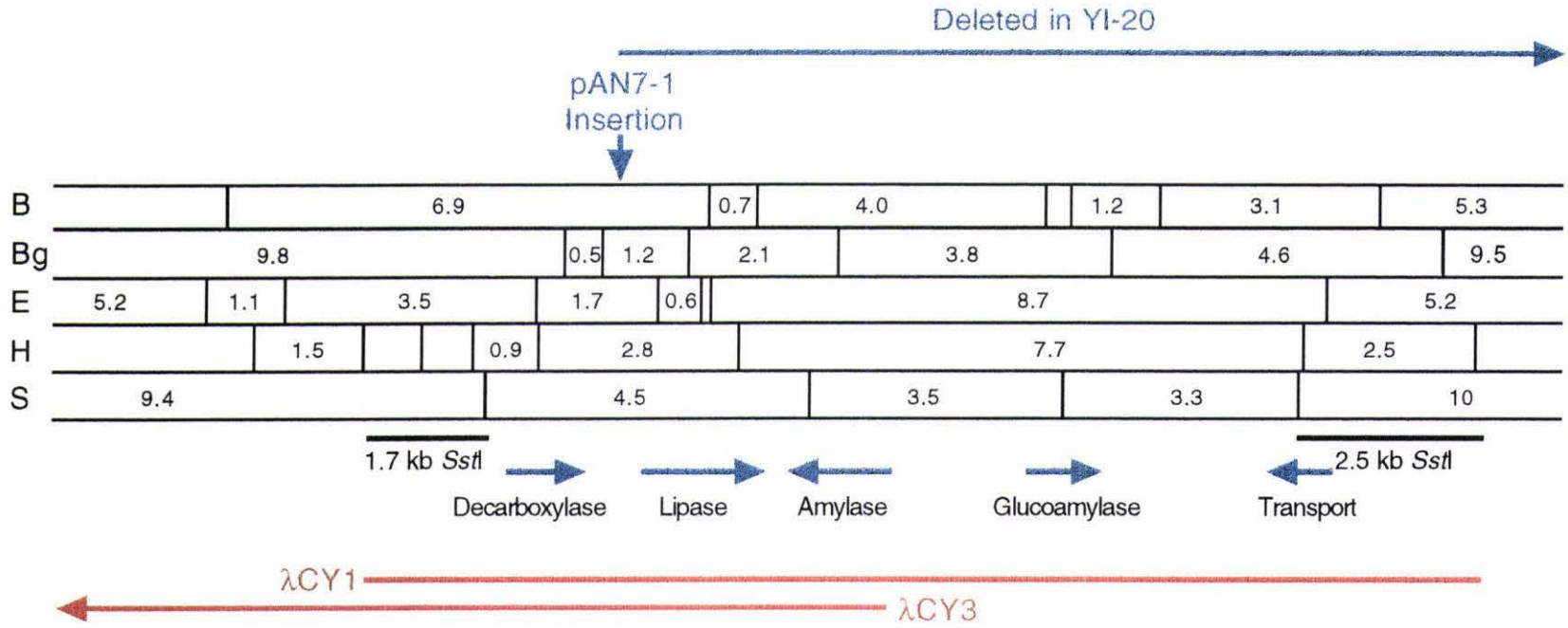



***Fig. 3.17*** *Restriction enzyme map of wild-type showing the point of pAN7-1 integration and the deleted region associated with pAN7-1 insertion.*

Abbreviations are B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; S, *Sst*I. Fragment sizes are shown in kilobases.

sequence similarity to a family of likely export proteins. If sequence similarities imply similarity of protein function, then these proteins are probably not involved in paxilline biosynthesis.

## 3.6 *Homologous Recombination Using a Replacement Construct*

To determine whether the Pax$^-$ phenotype of YI-20 was the result of pAN7-1 integration at the locus identified, gene disruptions were carried out using a replacement construct pPN1418 (Fig. 3.18 and 3.19). Four outcomes are possible following integration of this construct: (i) a single crossover into the 2.8 kb *Hin*dIII fragment, (ii) a single crossover into the 2.5 kb *Sst*I fragment (this is actually a 10.6 kb fragment in the *P. paxilli* genome but is truncated in λCY1), (iii) a double crossover between the 2.8 kb *Hin*dIII and the 2.5 kb *Sst*I fragments, and (iv) ectopic integration of the plasmid. It is also possible, as seen by Itoh and Scott (1994), that homologous recombination could generate new deletions.

### 3.6.1 *Construction of pPN1418 and Transformation into PN2013*

The plasmid pPN1418 was constructed in two steps (Fig. 3.18). The first step involved a three-way ligation (Section 2.13) between: (i) *Hin*dIII and *Sal*I linearised pUC118, (ii) the 2.8 kb *Hin*dIII fragment isolated from pPN1387, and (iii) a 2.3 kb *Hin*dIII/*Sal*I fragment containing the *hph* gene for hygromycin resistance from pCWHyg1 (Appendix A1.3). The resulting ligation was electroporated into *E. coli* XL1 cells (Section 2.13) and the transformants were screened for clones that had the 2.8 kb *Hin*dIII fragment oriented with the internal 0.5 kb *Bgl*II fragment closest to the pUC forward primer (pPN1414, Fig. 3.18). Plasmid pPN1414 was linearised with *Sst*I, CAP treated (Section 2.13), ligated with the 2.5 kb *Sst*I fragment isolated from pPN1406, and electroporated into XL1 cells (Section 2.13). These transformants were sereened for clones that had the 2.5 kb *Sst*I fragment oriented with the *Bgl*II site closest to the *Eco*RI site of pUC118 (pPN1418). This resulted in a plasmid, pPN1418, of 10.8 kb (Fig. 3.18 and 3.19).

Transformation of wild-type protoplasts with circular pPN1418 (Section 2.16) gave a total of 236 hygromycin resistant transformants per 5 µg of DNA. Of these, 123 were single-spore purified on PD plates containing 50 µg/ml of hygromycin (Section 2.2). Genomic DNA was isolated (Section 2.5) from these transformants and PCR used to determine the type of integration.

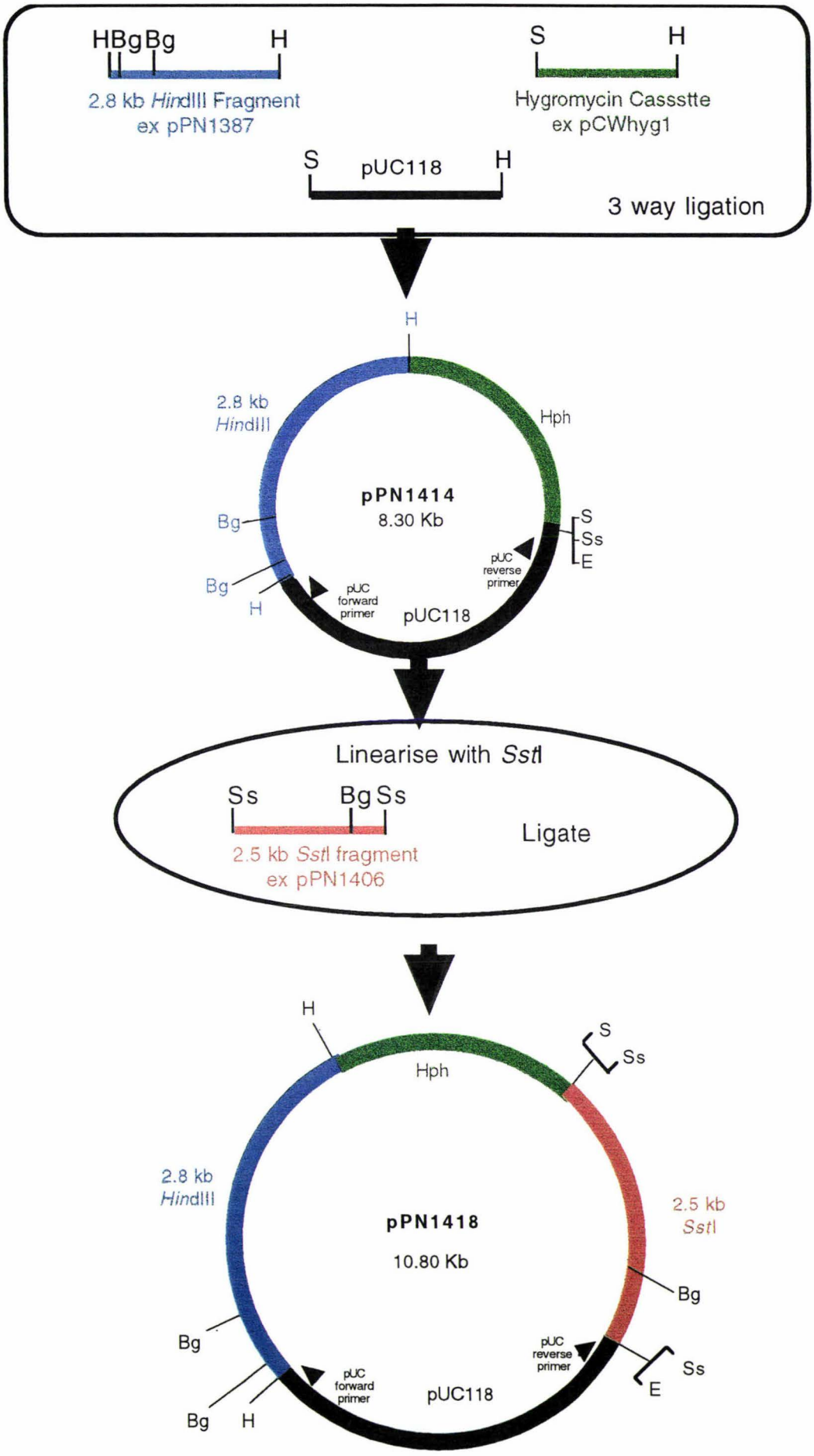



***Fig. 3.18*** *Construction of pPN1418*

Abbreviations are Bg, *Bgl*II; H, *Hin*dIII; S, *Sal*I; Ss, *Sst*I.

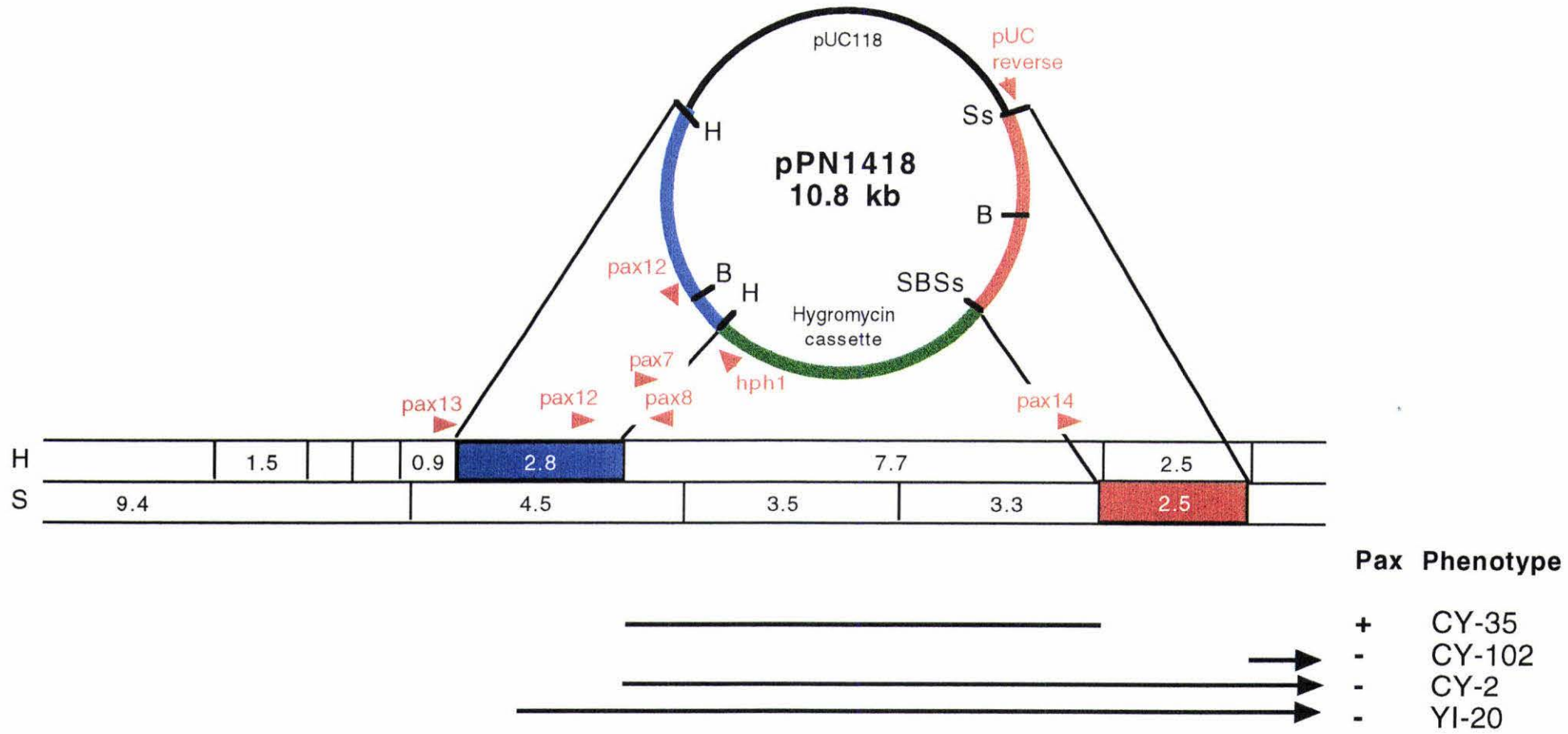


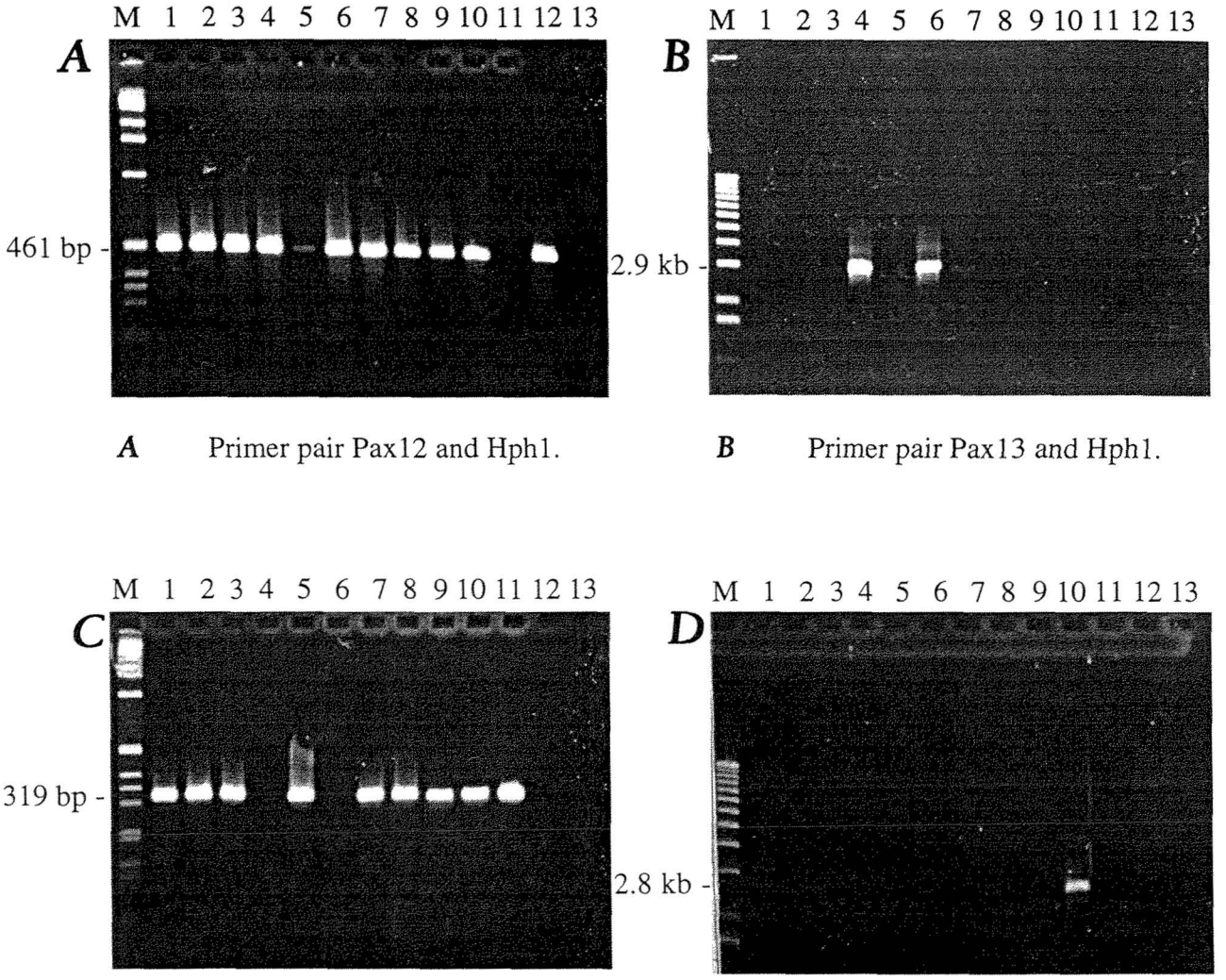

**Fig. 3.19** *Restriction map of wild-type and pPN1418 showing PCR strategy for determining the type of integration event.*

Abbrevations are: B, *Bam*HI, H, *Hin*dIII; S, *Sal*I; Ss, *Sst*I. Fragment sizes are shown in kilobases. The Black lines show the the deletion in each of the mutants.

Each of the 123 transformants was screened by PCR (Section 2.15) for homologous recombination (Fig. 3.19). To determine that the hygromycin resistance of these transformants was due to the integration of pPN1418, the DNA was amplified with the primer pair pax12 (a primer within the 2.8 kb *Hin*dIII fragment) and hph1 (a primer to the glucoamylase promoter attached to the hygromycin resistance gene). This resulted in the presence of a 461 bp fragment (Fig. 3.19 and 3.20A). To check for crossovers into the 2.8 kb *Hin*dIII fragment, DNA samples were amplified with pax13 (a primer within the 0.9 kb *Hin*dIII fragment) and hph1 (Fig. 3.19 and 3.20B). Ten of the transformants (Table 3.7 and 3.8) gave a 2.9 kb band indicative of this event. However, single crossovers that have associated deletions to the left of the 2.8 kb *Hin*dIII fragment will not be detected by this method. Two (1.6%) of the above transformants (Table 3.8), CY-35 and CY-39, that screened as single crossovers, were in fact, the result of a double crossover event. The absence of a 319 bp product with primers pax7 and pax8 (primers to the 7.7 kb *Hin*dIII fragment) is indicative of a deletion to the right-hand side, associated with homologous recombination into the 2.8 kb *Hin*dIII fragment (Fig. 3.19 and 3.20C). Three of the transformants lacked this product: two, CY-35 and CY-39, were deletions associated with a double crossover event, and the third, CY-2, contained a deletion that extended beyond the 2.5 kb (wild-type 10.6 kb) *Sst*I fragment. To check for a single crossover into the 2.5 kb *Sst*I fragment, DNA was amplified using the primers pax14 (a primer within the 3.3 kb *Sst*I fragment) and the pUC reverse primer (Fig. 3.19 and 3.20D); Ten (8.1%) produced the 2.8 kb band indicative of this event (Table 3.7 and 3.8). One, CY-102, was found by Southern analysis to contain a deletion beyond the region mapped. PCR analysis of the remaining transformants are described in Table 3.7. The remaining transformants (83.8%) that did not amplify PCR products with either the pax13/hph1 or pax14/rev primers were considered to be ectopic integrations. These transformants were not analysed for their ability to synthesis paxilline.

### 3.6.3 *Homologous Recombination Creates New Pax⁻ Mutants (With Deletions)*

All transformants determined as homologous crossovers were screened for their ability to produce paxilline by a competitive ELISA using a group-specific monoclonal antibody to paxilline (Section 2.17). Cultures that do not produce paxilline develop as a yellow colour, while those that have paxilline are clear or lighter. Two of the mutants, CY-2 and




*A* Primer pair Pax12 and Hph1.

*B* Primer pair Pax13 and Hph1.

*C* Primer pair Pax7 and Pax8.

*D* Primer pair Pax14 and pUC reverse.

**Fig. 3.20** *PCR analysis showing homologous integration events of pPN1418 transformants of Penicillium paxilli.*

***A - C*** PCR with the following DNA; CY-13 (lane 1), CY-14 (lane 2), CY-34 (lane 3), CY-35 (lane 4), CY-38 (lane 5), CY-39 (lane 6), CY-41 (lane 7), CY-43 (lane 8), CY-45 (lane 9), CY-51 (lane 10), wild-type (lane 11), pPN1418 (lane 12), $H_2O$ (lane 13). Lane M, 1 kb ladder.

***D*** PCR with the following DNA; CY-1 (lane 1), CY-2 (lane 2), CY-3 (lane 3), CY-4 (lane 4), CY-6 (lane 5), CY-8 (lane 6), CY-9 (lane 7), CY-10 (lane 8), CY-12 (lane 9), CY-13 (lane 10), wild-type (lane 11), YI-20 (lane 12), $H_2O$ (lane 13). Lane M, 1 kb ladder.

***Table 3.7*** *Summary of the homologous integration events of the pPN1418 transformants as shown by PCR and Southern analysis.*

| Transformant # | PCR primer pairs: Pax12 Hph1 | Pax13 Hph1 | Pax7 Pax8 | Pax14 Rev | Integration event | Paxilline phenotype |
|---|---|---|---|---|---|---|
| **CY-2** | √ | √ | **x** | **x** | **single *Hin*dIII Δ** | **negative** |
| CY-13 | √ | x | √ | √ | single *Sst*I | positive |
| CY18 | √ | √ | √ | x | single *Hin*dIII | positive |
| CY-26 | √ | x | √ | √ | single *Sst*I | positive |
| CY32 | √ | x | √ | √ | single *Sst*I | positive |
| CY-35 | √ | √ | x | x | double | positive |
| CY-39 | √ | √ | x | x | double | positive |
| CY-40 | √ | √ | √ | x | single *Hin*dIII | positive |
| CY-49 | √ | √ | √ | x | single *Hin*dIII | positive |
| CY-56 | √ | √ | √ | x | single *Hin*dIII | positive |
| CY-63 | √ | x | √ | √ | single *Sst*I | positive |
| CY-89 | √ | x | √ | √ | single *Sst*I | positive |
| CY-94 | √ | x | √ | √ | single *Sst*I | positive |
| CY-100 | √ | x | √ | √ | single *Sst*I | positive |
| CY-101 | √ | √ | √ | x | single *Hin*dIII | positive |
| **CY-102** | √ | **x** | √ | √ | **single *Sst*I Δ** | **negative** |
| CY-116 | √ | x | √ | √ | single *Sst*I | positive |
| CY-120 | √ | x | √ | √ | single *Sst*I | positive |
| CY-122 | √ | √ | √ | x | single *Hin*dIII | positive |
| CY-134 | √ | √ | √ | x | single *Hin*dIII | positive |

***Table 3.8*** *Analysis of homologous crossovers.*

| Type[a] | Site[b] | % | Ratio of Pax$^-$/Pax$^+$ |
|---|---|---|---|
| Single | *Hin*dIII | 6.5% | 1/8 |
| Single | *Sst*I | 8.1% | 1/10 |
| Double | *Hin*dIII/*Sst*I | 1.6% | 0/2 |
| Ectopic | ND[c] | 83.8% | ND[c] |

[a]The type of crossover was determined by PCR and Southern analysis. The percentage was determined from the 123 transformants analysed.
[b]The fragments involved in homologous recombination are those contained in the pPN1418 construct which contains the 2.8 kb *Hin*dIII and 2.5 kb *Sst*I fragments.
[c]Not determined.

CY-102, were found to be negative as can be seen by the strong yellow colour (Fig. 3.21). This result was confirmed by HPLC and TLC analysis (Section 2.17, Fig. 3.22 and 3.23). As with YI-20, no compounds with UV absorbance spectra similar to those of any known indole-diterpenoids were detected in the CY-2 and CY-102 extracts. Proposed intermediates for the paxilline biosynthetic pathway, including 13-desoxypaxilline, paspaline and paspaline B (Munday-Finch *et al.*, 1996) were also absent. TLC analysis, clearly shows that paspaline is missing from the three Pax$^-$ mutants. The compound eluting at 3.2 min in all fungal extracts was identified as ergosterol by comparison of its retention time and UV absorbance spectra with those of an authentic standard. All other transformants, including the double crossovers, CY-35 and CY-39, produced high levels of paxilline when assayed (Fig 3.22 and 3.23). The levels of paxilline found in CY-35 (3.7 mg/g dry weight of mycelia) were comparable to that of wild-type (3.4 mg/g dry weight of mycelia). These results confirm that the 7.6 kb region, between the 2.8 kb *Hin*dIII and 2.5 kb *Sst*I fragment is not required for paxilline biosynthesis.

The proposed recombination events based on the PCR results were confirmed by Southern analysis (Section 2.11) using genomic DNA from the following; wild-type, YI-20, CY-2, CY-102, CY-35, CY-39, CY-134 and CY-94. CY-2 and CY-102 were chosen because they screened as Pax$^-$. CY-35 and CY-39, both with deletions associated with them, were the result of a double crossover and are Pax$^+$. As all the single crossover transformants screened were Pax$^+$, only two, CY-134 and CY-94 (single crossovers into the 2.8 kb *Hin*dIII and 2.5 kb *Sst*I fragments, respectively), were selected for Southern analysis.

*Sal*I digests of genomic DNA from the strains stated above, probed with the pCWHyg1 (hygromycin resistance plasmid), are shown in Fig. 3.24. Wild-type (lane 1) does not contain the hygromycin cassette and, therefore, no bands hybridise. YI-20 (lane 2) containing a pAN7-1 tandem repeat, has the following hybridising fragments: (i) a 6.8 kb pAN7-1 repeat, (ii) a 10.8 kb band due to the inverted copy of pAN7-1, and (iii) a 4.2 kb band that contains left-hand flanking sequence. All of the pPN1418 transformants contain a band of 10.8 kb indicative of the integration of tandem repeats of the plasmid. This is likely due to transforming a circular plasmid. All but CY-39 (lane 7) have a single site of integration. CY-39 has approximately four points of integration, as can be seen by at least six other fragments hybridising with the probe. However, PCR and Southern results do indicate that CY-39 is due to a double crossover. Transformants that involved a homologous crossover into the 2.8 kb *Hin*dIII fragment, CY-2, CY-35, CY-39 and CY-134, (lanes 3, 6, 7 and 8) contained a common 7 kb band that was due to the left-hand flanking sequence. The right-hand flanking sequence corresponds to a 9.4 kb band

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Paxiline Standard Curve | 5000 | 3000 | 500 | 100 | 50 | 25 | 10 | 5 | 1 | 0.1 | B° | | A |
| | 5000 | 3000 | 500 | 100 | 50 | 25 | 10 | 5 | 1 | 0.1 | B° | | B |
| Duplicate Samples. Dilution A= Neat Dilution B= $^1/_{10}$ | 1A | 1A | 1B | 1B | 7A | 7A | 7B | 7B | 13A | 13A | 13 B | 13B | C |
| | 2A | 2A | 2B | 2B | 8A | 8A | 8B | 8B | 14A | 14A | 14 B | 14B | D |
| | 3A | 3A | 3B | 3B | 9A | 9A | 9B | 9B | 15A | 15A | 15 B | 15B | E |
| | 4A | 4A | 4B | 4B | 10A | 10A | 10B | 10B | | | | | F |
| | 5A | 5A | 5B | 5B | 11A | 11A | 11B | 11B | | | | | G |
| | 6A | 6A | 6B | 6B | 12A | 12A | 12B | 12B | | | | | H |

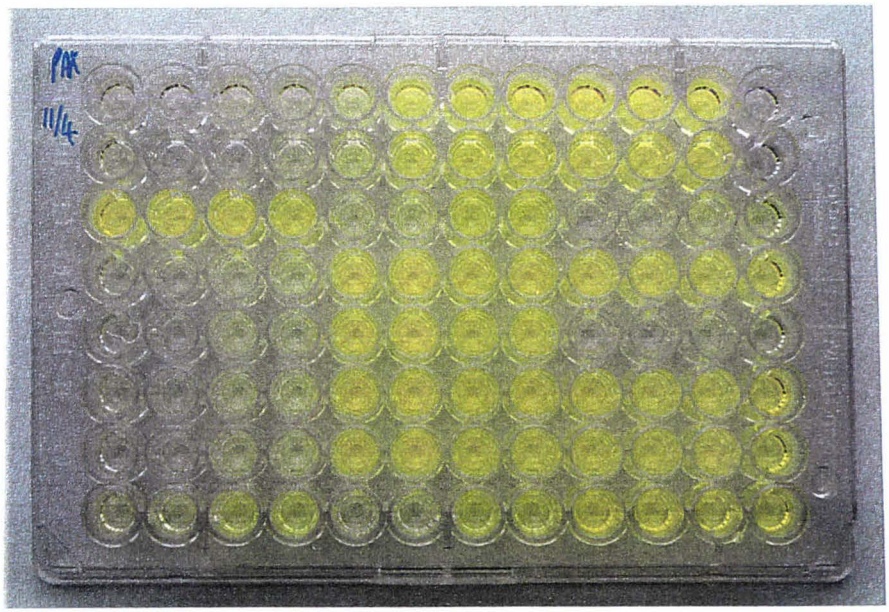


***Fig 3.21*** *Monoclonal antibody-based ELISA assay used for screening for paxilline phenotype.*

With reference to the numbering template at the top, culture filtrates of samples were analysed as follows: CY-2 (wells 1); CY-13 (wells 2 and 3); CY-26 (wells 4 and 5); CY-94 (wells 6 and 7); CY-102 (wells 8 and 9); YI-20 (wells 10, 11 and 14); CY-116 (wells 12 and 13); wild-type (wells 15).

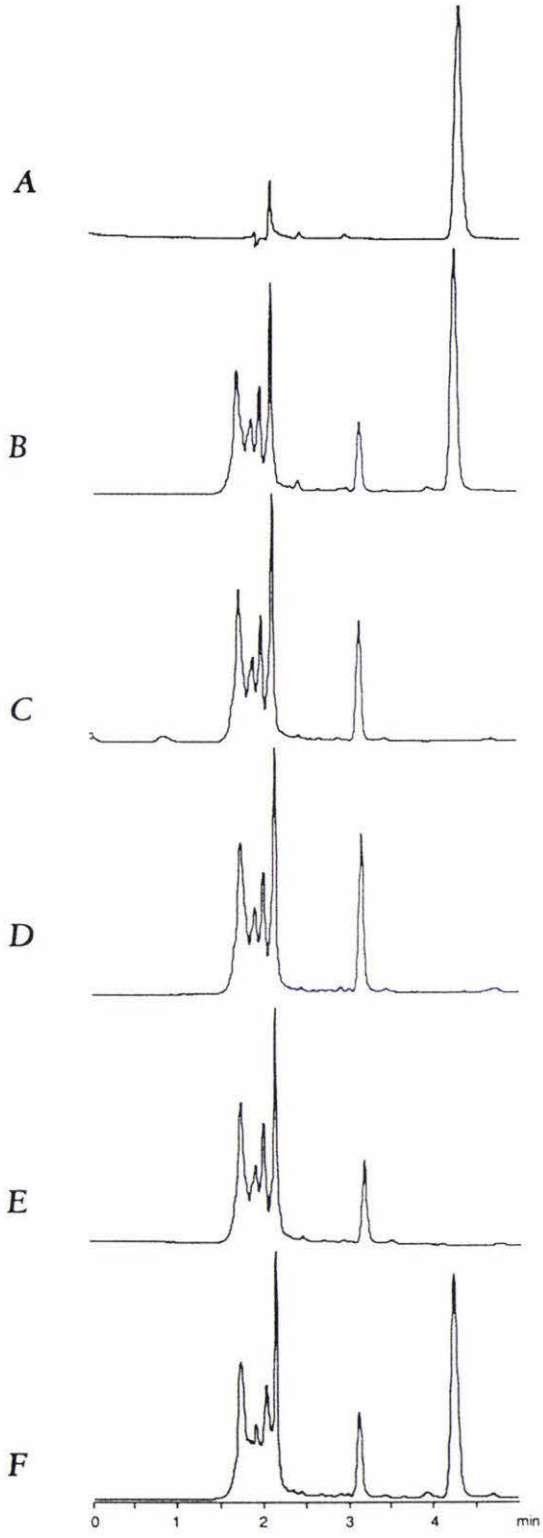



**Fig. 3.22** *HPLC analysis confirming the paxilline phenotype of the deletion mutants.*

| | |
|---|---|
| ***A*** | Authentic paxilline |
| ***B*** | Wild-type |
| ***C*** | YI-20 |
| ***D*** | CY-2 |
| ***E*** | CY-102 |
| ***F*** | CY-35 |

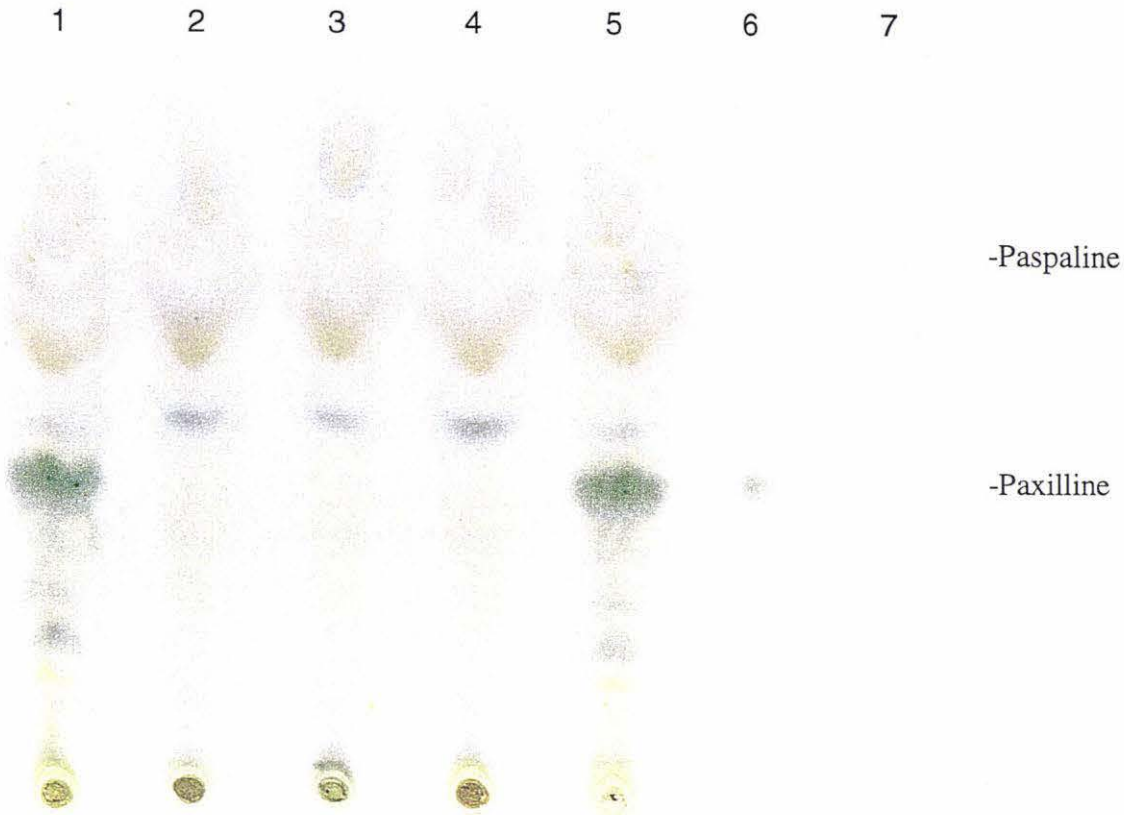



**Fig. 3.23** *TLC analysis of wild-type and the plasmid generated mutants of P. paxilli.*

Samples (15 µl) were spotted as follows: Wild-type (lane 1); YI-20 (lane 2); CY-2 (lane 3); CY-102 (lane 4); CY-35 (lane 5); authentic paxilline (lane 6); authentic paspaline (lane 7).

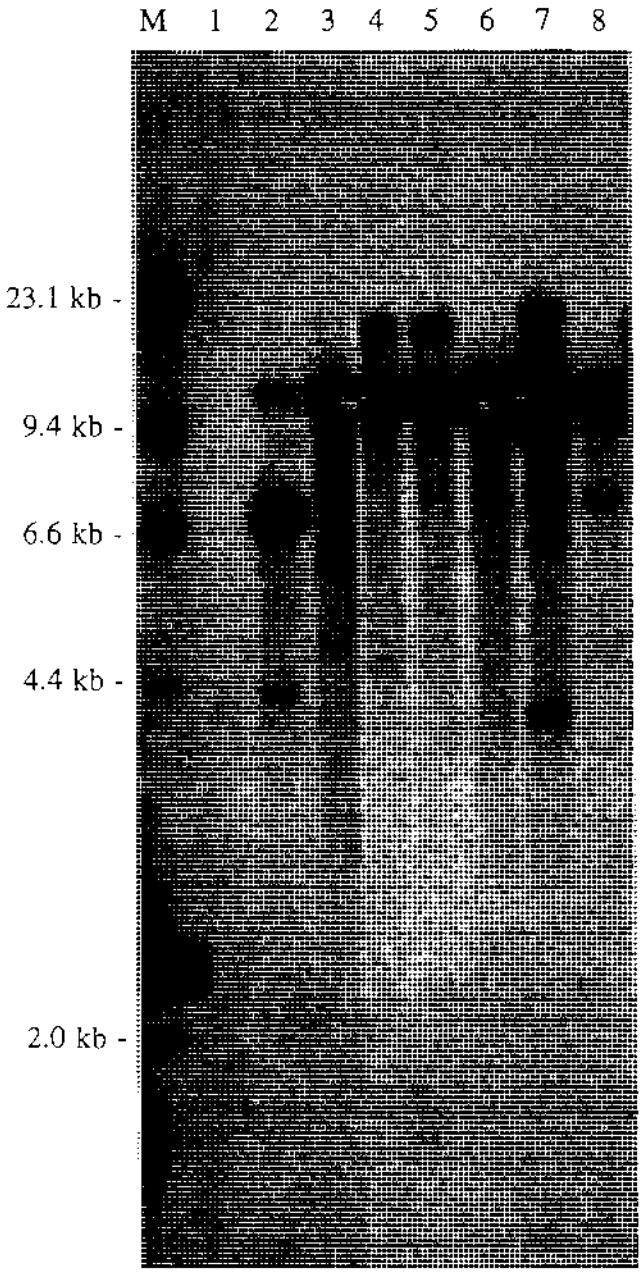



**Fig. 3.24** *Mapping the integration events of pPN1418.*

Southern analysis of 1 μg of *Sal*I digested genomic DNA from wild-type (lane 1), YI-20 (lane 2) and the following transformants; CY-2 (lane 3), CY-102 (lane 4), CY-94 (lane 5), CY-35 (lane 6), CY-39 (lane 7), CY-134 (lane 8), probed with [α-$^{32}$P]dCTP labelled pCWHyg1. Lane M, λ *Hin*dIII.

for CY-134 (a single crossover). CY-2 (lane 3) has a smaller right-hand fragment of 5.8 kb. No right-hand band is expected to be seen in CY-35 (lane 6) as only the tandem repeat and left-hand flanking sequences can hybridise. The difference between CY-2 (lane 3) and CY-134 (lane 8) with regards to the size of their right-hand fragment confirms the earlier PCR result that indicated a deletion has occurred in CY-2. CY-102 (lane 4) and CY-94 (lane 5) have a common hybridising left-hand fragment of approximately 15.1 kb. A very faint band of 4.4 kb, which was detected in CY-102, is attributed to a right-hand flanking sequence and is likely due to the crossing out of pPN1418 part way through the pUC118 part of the plasmid. As no such band is detected in CY-94, this confirms that CY-102 has a deletion associated with the homologous recombination of the plasmid.

Hybridisation results of *Bam*HI digested DNA of the same samples as shown in Fig. 3.24, probed with the 2.5 kb *Sst*I fragment (Section 2.11) are as follows (Fig. 3.25): wild-type (lane 1) has two hybridising fragments of the expected sizes of 3.1 kb and 5.3 kb, YI-20 (lane 2) is deleted beyond the 2.5 kb band and therefore does not hybridise. All of the pPN1418 transformants contain the expected repeat bands of 6.8 kb and 1.3 kb. CY-102 (lane 4) and CY-94 (lane 5) also contain the 3.1 kb wild-type band. CY-2 (lane 3) has a 4.4 kb band due to right hand flanking sequence and the left hand side can not hybridise. CY-35, CY-39, CY-134 and CY-94 (lanes 6, 7, 8 and 5) all contain the wild-type 5.3 kb band indicating that no deletion has occurred. CY-102 (lane 4) does not contain the expected 5.3 kb right hand band as it has a deletion beyond the 2.5 kb S*st*I fragment.

## 3.7 *Mapping of CY-2 for Plasmid Rescue and IPCR*

Data from the *Sal*I digested transformants hybridising to pCWhyg1 (Fig. 3.24) shows CY-2 has a right-hand flanking band of 5.8 kb hybridising to approximately the same intensity as that expected of the left-hand flanking band (7 kb). As these left- and right-hand fragments hybridise to the same intensity, this suggests that a fragment of approximately 2.3 kb of pUC118 (the same size as the hybridising hygromycin cassette of pCWHyg1) remains adjacent to the right-hand flanking sequence. This 5.8 kb *Sal*I band must be due to the crossing out of pPN1418 part way through pUC118. This was confirmed by hybridisation of a *Sal*I digest of CY-2 genomic DNA probed with pUC118. Southern analysis of CY-2 using the enzymes *Bam*HI, *Eco*0109I, *Eco*RI, *Hin*dIII, *Sal*I, *Sst*I and the double digests *Eco*0109I/*Eco*RI and *Eco*RI/*Hin*dIII probed with pUC118 (Fig. 3.26) confirmed a partial copy of the pUC118 vector remained adjacent to the right

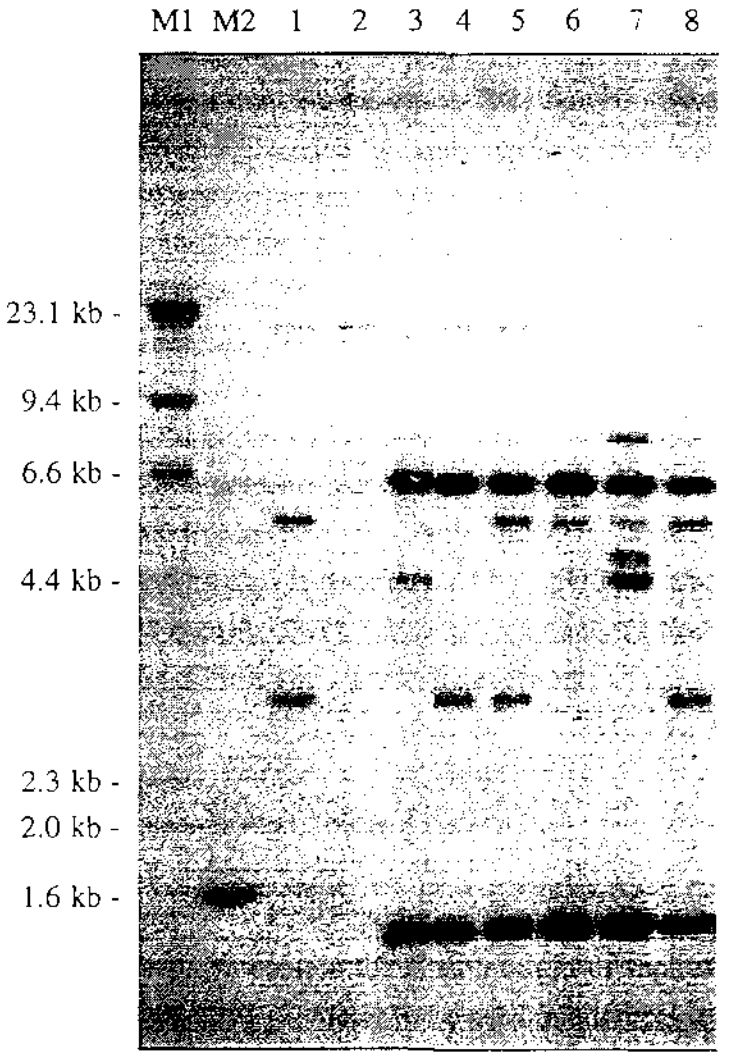



*Fig. 3.25* *Mapping the integration events of pPN1418.*

Southern analysis of 1 µg of *Bam*HI digested genomic DNA from wild-type (lane 1), YI-20 (lane 2) and the following transformants; CY-2 (lane 3), CY-102 (lane 4), CY-94 (lane 5), CY-35 (lane 6), CY-39 (lane 7), CY-134 (lane 8), probed with [α-$^{32}$P]dCTP-labelled 2.5 kb *Sst*I fragment from pPN1406. Lane M1, λ *Hin*dIII. Lane M2, 1 kb ladder.

hand flanking sequence (Fig. 3.27). As confirmed previously, the *Sal*I digest (lane 7) shows the expected 10.8 kb tandem repeat band and 5.8 kb right-hand flanking sequence. Of the 5.8 kb fragment, 4.8 kb consists of the 2.5 kb *Sst*I and 2.3 kb pUC118 fragments with the remaining 1 kb due to right-hand flanking sequence. The *Bam*HI digest (lane 1) shows a 6.8 kb repeat band and the 4.4 kb right-hand flanking sequence band (also seen in Fig. 3.25). Of this 4.4 kb band, 1.1 kb is due to part of the 2.5 kb *Sst*I fragment, so the remaining sequence must consist of ~2.3 kb of pUC118 and ~1.0 kb of right-hand flanking sequence. The *Eco*RI digest (lane 4) has a 4.5 kb tandem repeat band and a 3.4 kb band which would consist of ~1.1 kb of right-hand flanking sequence. The *Sst*I digest has a 8.2 kb tandem repeat band and a 3.4 kb band that would contain ~1.1 kb of right-hand flanking sequence. The *Hin*dIII digest only has a 5.6 kb fragment which would be due to the tandem repeat. However, it is possible that this contains the single copy right-hand fragment (as no other hybridising fragment relating to this fragment is seen). Such a band would contain ~2.3 kb of the 2.5 kb *Sst*I fragment, 2.3 kb of pUC118 and ~1.0 kb of right-hand flanking sequence. It appears from this Southern analysis that the five separate restriction digests analysed map to the same point in the genome (Fig. 3.27)

### 3.7.1 *Attempts to Rescue the Right-Hand Side of CY-2*

The 2.3 kb partial copy of pUC118 would contain the *E. coli* origin of replication and enough of the ampicillin resistance gene and promoter to be used in attempts to plasmid rescue the right-hand side flanking the deletion in CY-2 with self-ligated *Bam*HI digested DNA (Section 2.13). Of the 252 resulting plasmid-rescued transformants, 58 were screened using *Bam*HI digested DNA prepared by the rapid boil method (Section 2.5). All the rescued clones contained the 6.8 kb band due to the tandem repeat and not the 4.4 kb band expected for a clone containing right-hand flanking sequence. Although there is enough pUC118 present for such a rescued clone to contain both ampicillin resistance and the origin of replication required for propagation in *E. coli*, the only sequence rescued in *Bam*HI digested CY-2 spanned the tandem repeat.

Inverse PCR (IPCR) was also used (Section 2.15) with *Bam*HI and *Eco*RI digested/self ligated CY-2 genomic DNA. PCR with this self-ligated DNA using the pUC reverse and pUCap primers (Fig. 3.27) resulted only in the product associated with the presence of the tandem repeat of pPN1418 and not with the right-hand flanking sequence (data not shown).

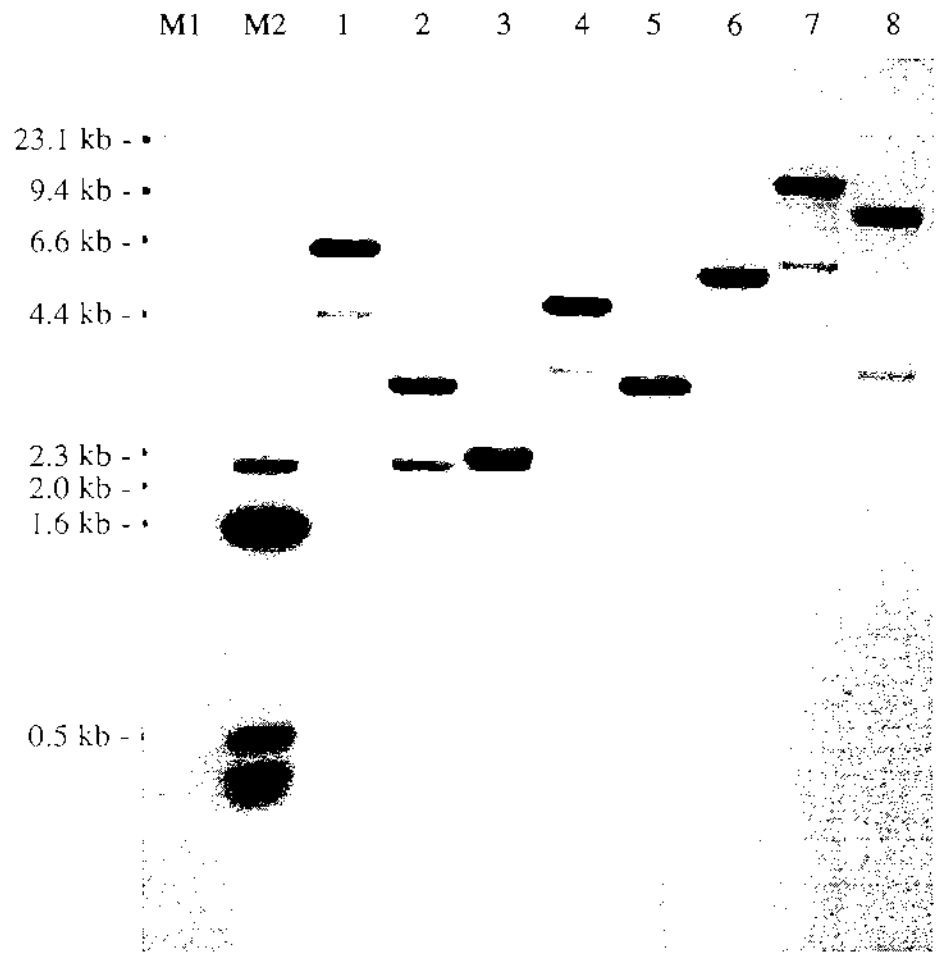



**Fig. 3.26** *Southern analysis of CY-2 showing right-hand flanking sequence.*

Southern analysis of 1 µg of CY-2 DNA digested with *Bam*HI (lane 1), *Eco*01091 (lane 2), *Eco*01091/*Eco*RI (lane 3), *Eco*RI (lane 4), *Eco*RI/*Hin*dIII (lane 5), *Hin*dIII (lane 6), *Sal*I (lane 7), *Sst*I (lane 8) hybridised to [α-$^{32}$P]dCTP-labelled pUC118. Lane M1, λ *Hin*dIII. Lane M2, 1 kb ladder.

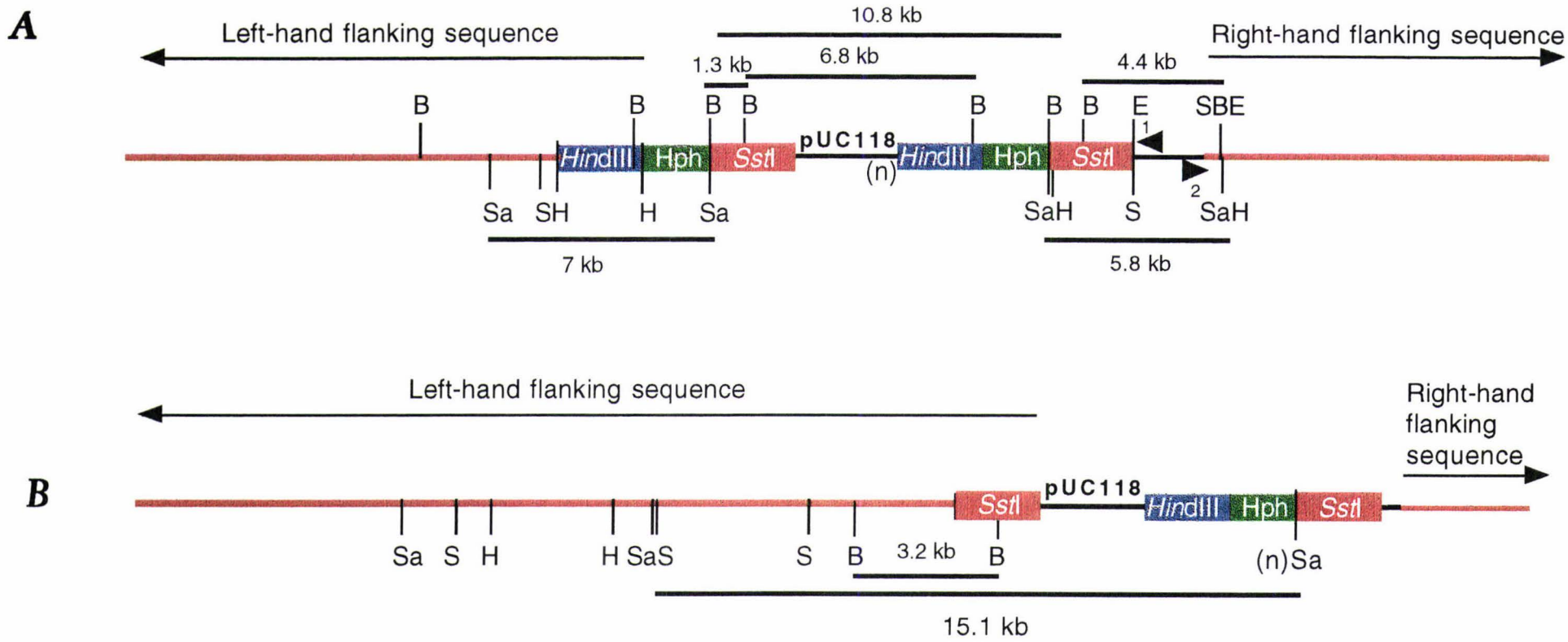



***Fig. 3.27*** *Restriction enzyme maps of CY-2 and CY-102*

***A*** Restriction enzyme map of CY-2.

***B*** Restriction enzyme map of CY-102.

Abbreviations: B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; Sa, *Sal*I; S, *Sst*I; 1, pUC reverse primer; 2 pUCap primer. Red lines are genomic flanking sequence. Selected restriction enzyme sites are shown.

CY-102 only contains a very small amount of pUC118 adjacent to the right-hand flanking sequence and therefore no attempts have been made to rescue this sequence or analyse it further.

## 3.8 *Characterisation of Karyotypes*

As the parental strain PN2013 was not genetically marked, RAPD analysis (Fig. 3.28 A, B and C) was used to confirm that the Pax$^-$ mutants (YI-20, CY-2 and CY-102), and the double crossover (CY-35) were derived from the wild-type strain (PN2013). A unique RAPD banding pattern was generated for each primer (GT02, RC05 and RC09) and this pattern was conserved between wild-type and all mutants, confirming that they were derived from PN2013.

The karyotype of each of the above strains was also determined (Fig. 3.29) using conditions previously established to resolve chromosomes of wild-type *P. paxilli* (Itoh *et al.* 1994). Under these conditions the chromosomal banding of the mutants was identical to wild-type (Fig. 3.29). However, when this gel was blotted and probed with the 0.5 kb *Bgl*II fragment (ex pPN1375), which is adjacent to the site of pAN7-1 insertion in YI-20 (Fig. 3.7), chromosomal band V hybridised in wild-type, CY-2, CY-102 and CY-35 (Fig. 3.30A, lanes 1, 3, 4 and 5) while chromosomal band VI hybridised in YI-20 (lane 2). This result would indicate that not only are sequences deleted in YI-20 but there has also been a chromosomal rearrangement in this strain such that a band of the same mobility as wild-type chromosome V is retained and a new band of about the same mobility as wild-type chromosome VI is generated. In contrast, insertion of the replacement construct pPN1418 in mutants CY-2 and CY-102 has not resulted in a chromosomal rearrangement. The hybridising chromosome V of CY-2, CY-102 and CY-35 (lanes 3, 4 and 5) is more intense as the 0.5 kb *Bgl*II fragment is present in the integrating plasmid, pPN1418, whereas in wild-type and YI-20 this fragment is present as a single copy. Further analysis of this CHEF gel (Fig. 3.30B) shows that a 1 kb *Hin*dIII/*Kpn*I probe, corresponding to sequences that flank a pAN7-1 mutant (YI-186) generated by restriction enzyme mediated (REM) integration (Itoh and Scott, 1997), hybridised to a band of the same mobility as chromosome VI in wild-type and all mutants (Fig. 3.30B). These two results would suggest that the band labelled as chromosome V in wild-type (Fig. 3.29A, lane 1) consists of two chromosomes of similar mobility and that one of these is rearranged in YI-20. In an attempt to resolve these large chromosomes, electrophoresis conditions were modified to those used recently to resolve the chromosomes of *A. nidulans* (Geiser *et al.*, 1996). Under these conditions (Fig. 3.31) chromosomal band V of wild-type (Fig. 3.29, lane 1)

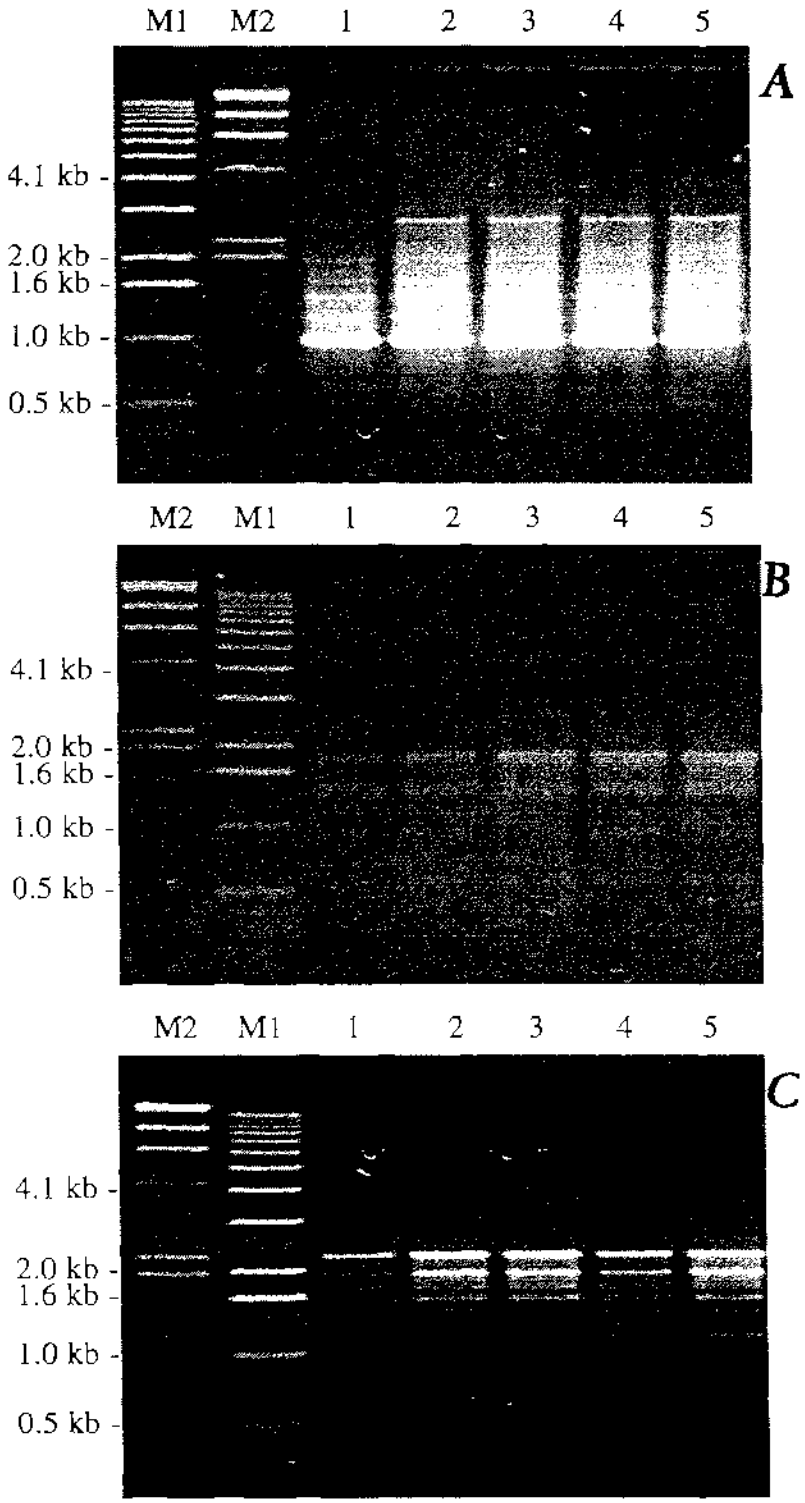



***Fig. 3.28*** *RAPD analysis of wild-type, YI-20 and pPN1418 transformants.*

RAPD analysis of wild-type (lane 1), YI-20 (lane 2), CY-2 (lane 3), CY-102 (lane 4), CY-35 (lane 5). Lane M1, 1 kb ladder. Lane M2, λ *Hin*dIII.

***A*** Using RAPD primer GT02.

***B*** Using RAPD primer RC05.

***C*** Using RAPD primer RC09.

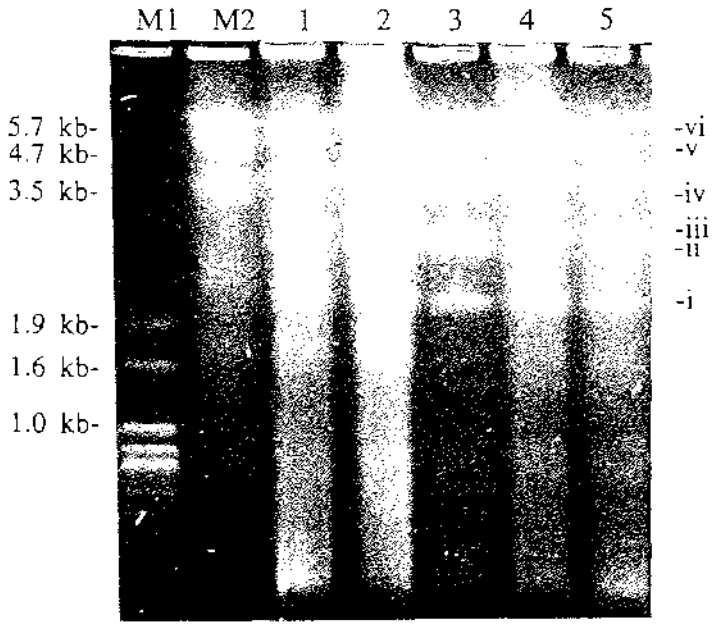



**Fig. 3.29** *CHEF analysis of chromosomal DNA of wild-type and the plasmid generated mutants of P. paxilli.*

Wild-type (lane 1), YI-20 (lane 2), CY-2 (lane 3), CY-102 (lane 4), CY-35 (lane 5). Lane M1, *S. cerevisiae*. Lane M2, *S. pombe*. Using the conditions of Itoh *et al.* (1994), 0.6% (w/v) chromosomal grade agarose in 0.5 x TBE at 60 volts with the following switching times; 500 sec for 25 h, 1000 sec for 30 h, 1500 sec for 50 h and 2500 sec for 32 h.

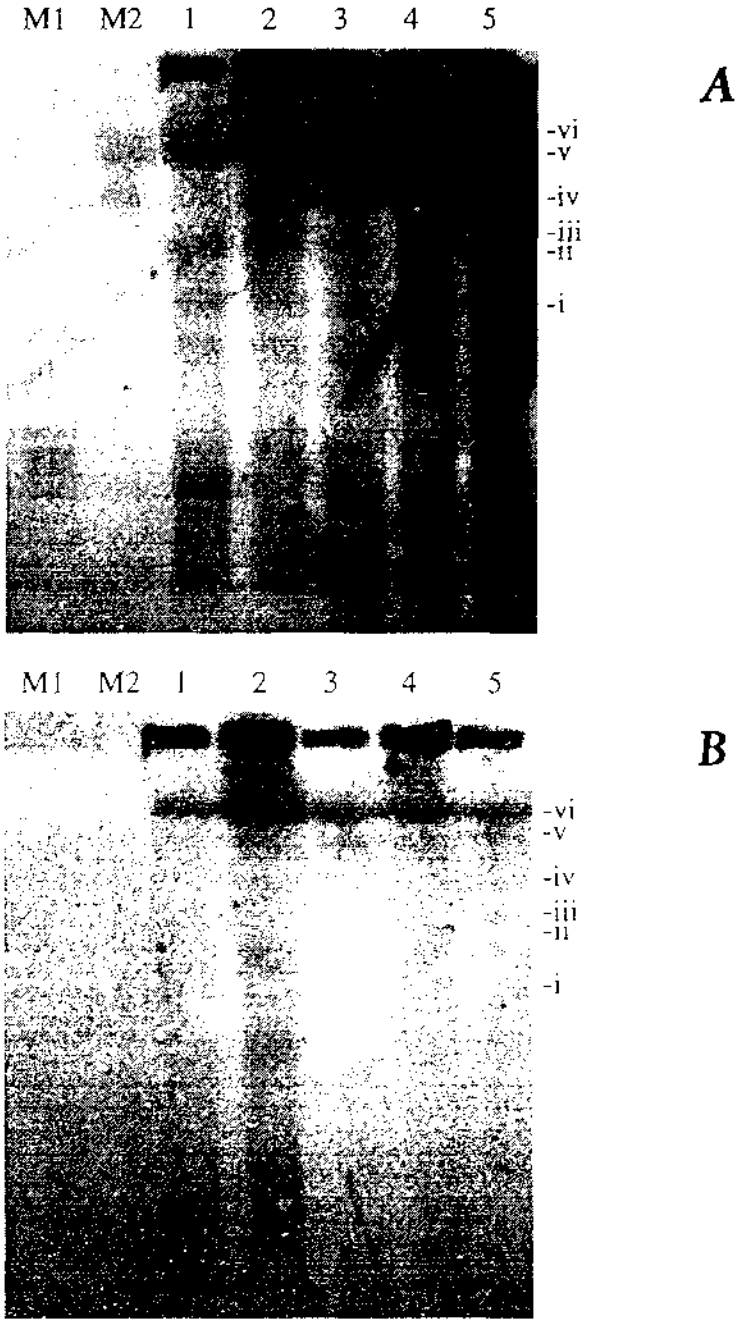



**Fig. 3.30** *CHEF analysis of chromosomal DNA of wild-type and the plasmid generated mutants of P. paxilli.*

Southern analysis of Fig. 3.29. Wild-type (lane 1), YI-20 (lane 2), CY-2 (lane 3), CY-102 (lane 4), CY-35 (lane 5). Lane M1, *S. cerevisiae*. Lane M2, *S. pombe*.

***A*** Probed with 0.5 kb *Bgl*II fragment.

***B*** Probed with 1 kb *Hin*dIII/*Kpn*I fragment.

is resolved into two bands, designated as Va and Vb (Fig. 3.31A and 3.31B). Band Va is absent from YI-20 and a new band is apparent immediately below band VI (lane 2). Mutants CY-2 (lane 3) and CY-102 (lane 4) have a band of slightly greater mobility than Va, consistent with a deletion of 120 - 150 kb. Mutant CY-35 appears to lack Va (lane 5), but, as it consists of a small deletion and multiple copies of the integrating pPN1418 plasmid, the Va chromosome is expected to be slightly larger than normal. When a blot of this gel was probed with the 1.7 kb *Sst*I end-fragment of λCY1 (Fig 3.17), which is adjacent to the site of homologous insertion of pPN1418 but not contained within the plasmid (Fig. 3.19), chromosome Va (lane 1) or the modified derivative of this band (lanes 2 - 5) hybridised (Fig. 3.32A). Reprobing of this blot with pAN7-1 confirmed that the same chromosome Va bands that hybridised with flanking sequence also hybridised with pAN7-1 (Fig. 3.32B, lanes 2 - 5).

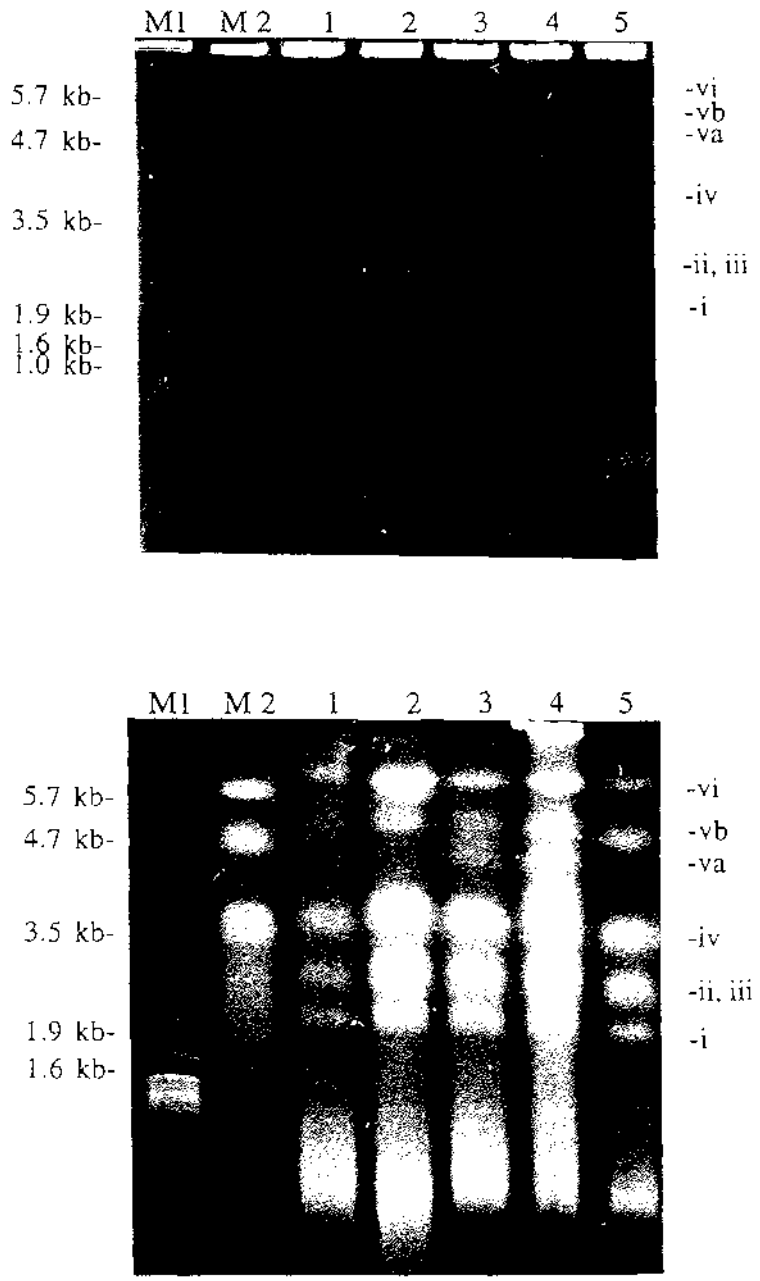



**Fig. 3.31** *CHEF analysis of chromosomal DNA of wild-type and the plasmid generated mutants of P. paxilli.*

Ethidium bromide stained CHEF gels of Wild-type (lane 1), YI-20 (lane 2), CY-2 (lane 3), CY-102 (lane 4), CY-35 (lane 5). Lane M1, *S. cereviseae*. Lane M2, *S. pombe*. Using modified conditions of Geiser *et al.* (1996), 0.6% (w/v) chromosomal grade agarose in 0.5 x TBE at 47 volts with the following switch times; 3000 sec for 73 h, 2700 sec for 18 h, and 2200 sec for 73 h.

***A*** Half a plug used.

***B*** A whole plug used.

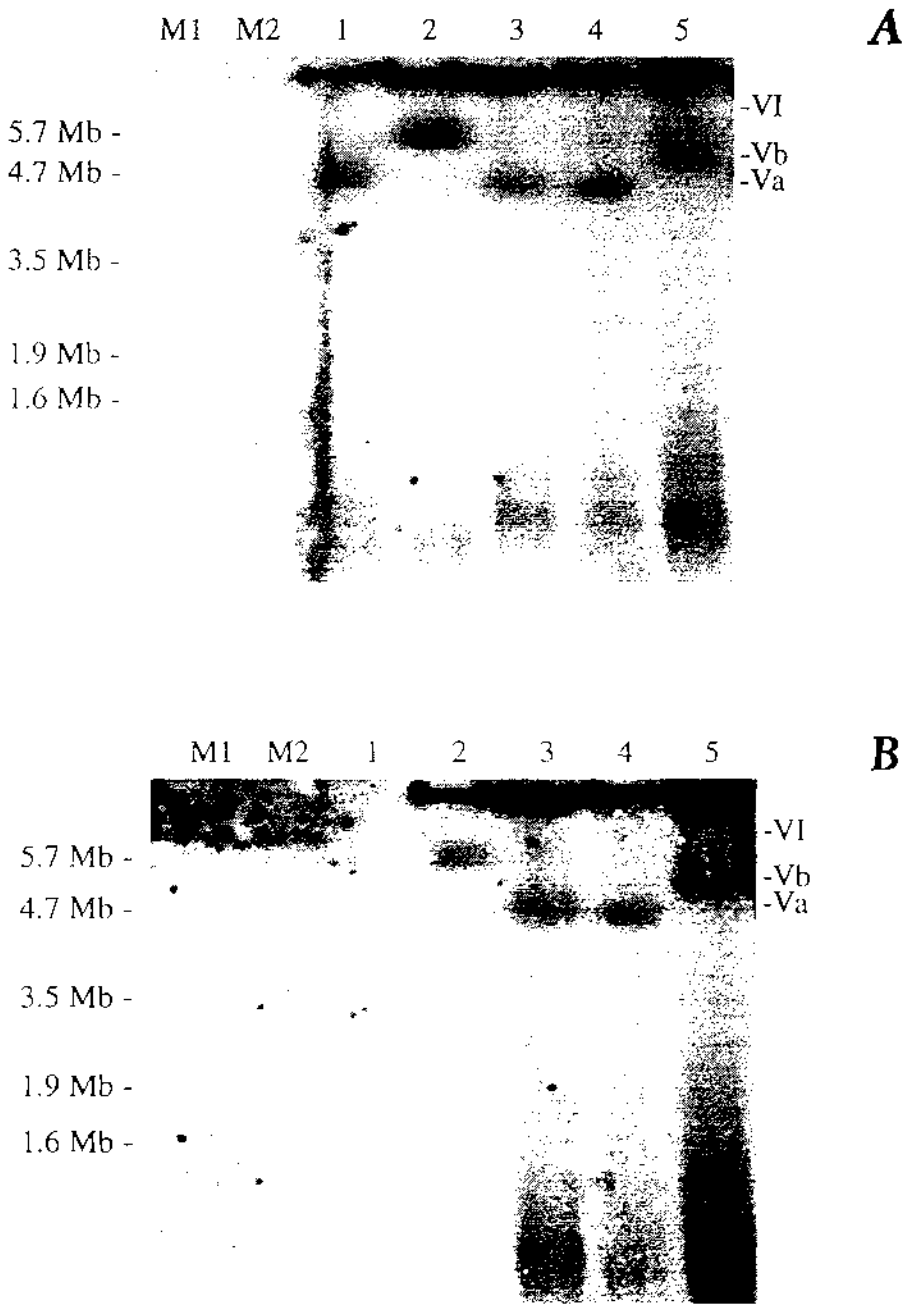



**Fig. 3.32** *Southern analysis of chromosomal DNA wild-type and plasmid-generated mutants of P. paxilli.*

A blotted CHEF gel (Fig3.26A) wild-type (lane 1), YI-20 (lane 2), CY-2 (lane 3), CY-102 (lane 4), CY-35 (lane 5). Lane M1, *S. cereviseae*. Lane M2, *S. pombe*.

***A*** Probed with the 1.7 kb *Sst*I fragment.

***B*** Probed with pAN7-1.

# *Chapter Four*
# *Discussion*

This work was initiated by Itoh (Section 1.6), where 600 pAN7-1-tagged *P. paxilli* transformants were analysed for their ability to produce paxilline in submerged culture. Of those 600 transformants one, YI-20, did not produce paxilline as detected by competitive ELISA (Itoh, unpublished data). The monoclonal-antibody (M-03/01) used in this assay recognises both paxilline and an array of compounds structurally related to paxilline (Miles *et al.*, 1995). Therefore, this negative result indicates that the paxilline biosynthetic pathway in YI-20 is blocked at such an early stage that cross-reacting indole-diterpenoids are not present. This phenotype was confirmed by HPLC and TLC analysis where paspaline, which is an early key intermediate of paxilline biosynthesis and which has a poor cross-reactivity to the antibody, was not detected (Section 3.6.3). The HPLC and TLC analysis also confirmed that no other known paxilline intermediates or new indole-diterpenoid compounds were produced. The broad specificity of the M-03/01 antibody may have resulted in the selection of a mutant that is blocked at an early stage of paxilline biosynthesis.

Characterisation of YI-20 shows that pAN7-1 has integrated at one site, but exists as a tandem repeat at this location (Section 1.6). Although linearised pAN7-1 was used in the transformation, thereby reducing the chances of obtaining tandem repeats, these repeats were still a major outcome of integration in *P. paxilli*. A study of pAN7-1 integrations in *P. paxilli* using a linearised plasmid has shown that only 37% of transformants had a single copy of pAN7-1 integrated at a single site and that the remaining 63% of transformants had tandem repeats of the plasmid (Itoh, 1994). In contrast, 82% of transformants where plasmids were integrated in *P. paxilli* in the presence of a restriction enzyme contain a single copy of the plasmid (Itoh and Scott, 1997).

Southern analysis of YI-20 showed that the pAN7-1 tandem repeat contained at least 5 copies of the plasmid (Section 3.2) and was further complicated by the presence of a head-to-head inverted repeat (Section 3.1.2). Detection of fragments, that contain wild-type sequences flanking the point of pAN7-1 integration is difficult, as pAN7-1 hybridises strongly to the multi-copy pAN7-1 bands and many of these fragments are the same size as those containing the single-copy flanking sequence (Section 3.1). However, a map of the left-hand integration site was determined (Fig. 3.1).

Plasmid rescue (Fig. 3.7) was used to isolate the genomic sequence adjacent to the left-hand side of the integrating plasmid (Section 3.3). Despite many attempts, I was not able to either detect or isolate the genomic sequences that flank the right-hand side of the plasmid (Section 3.1.2 and 3.4). As I was unable to identify a fragment that contains right-hand flanking sequence, this makes it difficult to design an experiment to rescue this sequence. The difficulty of obtaining this sequence could be due to complications

associated with the tandem repeat or that a partial copy of pAN7-1 flanks the right-hand side and a suitable probe (e.g. pUC118) was not used to detect this.

Analysis of sequence data surrounding the left-hand integration site of the plasmid showed that this integration event did not occur at a site within the genome that has sequence similarity to pAN7-1 (Section 3.5.3). However, a complete analysis of pAN7-1 integration in *P. paxilli* can not be made until we can compare both sides of the target site and a number of different pAN7-1 integration events. The loss of the *Bam*HI site adjacent to the genomic flanking sequence has been attributed to the plasmid having an 8 bp deletion from this site (Section 3.5.3). It is also expected that the right-hand side of the tandem repeat does not contain a complete copy of pAN7-1. Analysis of junction sequences associated with heterologous plasmid integration in other filamentous fungi shows little homology between the incoming plasmid and the target sites, and has also shown truncation of the integrating plasmid (Razanamparany and Begueret, 1988; Asch *et al.*, 1992).

Genomic sequences surrounding the point of pAN7-1 integration were isolated from a wild-type λGEM-11 library (Section 3.5). Two overlapping λ clones, λCY1 and λCY3, span ~20 kb across the integration site. Hybridisation of λCY1 against wild-type and YI-20 genomic DNA detected a large deletion in YI-20 that extended beyond the λCY1 boundary (Section 3.5.4). From this data we know the deletion is of a size greater than 12 kb. Deletions are a common outcome of pAN7-1 integration into the genome of *P. paxilli*. Of five mutants with a brown spore phenotype ($Brs^-$), four contained plasmid integrations that were at different sites within a 20 kb distance, and all five $Brs^-$ had deletions associated with them at this locus (Itoh and Scott, 1994).

DNA sequence obtained across the deleted portion contained in λCY1 revealed a number of open reading frames (ORFs). These identified ORFs had high similarity to a phenylacrylic acid decarboxylase, a lipase, an α-amylase, a glucoamylase and a transport protein. These proteins are unlikely to be involved in paxilline biosynthesis (Section 3.5.2 and 3.5.4). As there are no detectable differences between the growth of wild-type and YI-20, the deleted sequences are dispensable. Deletions have been shown to be a common outcome of heterologous plasmid integration in filamentous fungi (Kistler and Benny, 1992; Itoh and Scott, 1994).

Pulse-field gel electrophoresis of YI-20 revealed that a chromosome rearrangement has occurred upon integration of pAN7-1 (Section 3.8). Comparison of the chromosome banding patterns between YI-20 and wild-type (Fig. 3.30) showed that YI-20 lacked chromosome Va and instead has a chromosome that is increased in size by about 1 Mb.

Such an increase in size can not be explained by the presence of multiple copies of the integrating plasmid, as the expected size increase of this chromosome with 5 copies of pAN7-1, would only be 35 kb and YI-20 is known to have a deletion of at least 12 kb (Section 3.5.4). Southern analysis revealed that the rearranged chromosome did in fact contain not only the integrating plasmid but also DNA known to be contained on chromosome Va in wild-type (Section 3.8). With limited genetic markers available it is not possible to determine whether the rearrangement observed was the result of a reciprocal or non-reciprocal translocation. Chromosome deletions and rearrangements are a frequent outcome when ectopic integration events have occurred (Wernars *et al.*, 1986; Asch *et al.*, 1992; Kistler and Benny, 1992; Perkins *et al.*, 1993; Itoh and Scott, 1994). Rearrangements were found in 10% of transformants of *Neurospora crassa*, where in 6 of the 10 translocations examined the transforming plasmid was genetically linked to the breakpoints (Perkins *et al.*, 1993). In contrast, untransformed protoplasts of *Neurospora crassa* contained almost no rearrangements. Rearrangements have also been found in other secondary metabolite biosynthetic pathways as an outcome of different mutagenesis procedures, including ultraviolet (UV) irradiation (Hodges *et al.*, 1994) and chemical treatments (Fierro *et al.*, 1996).

Due to the genetic complexity of YI-20, in which there are multiple copies of the integrating plasmid, a deletion and a chromosome rearrangement, it was important to establish whether the Pax$^-$ phenotype was the result of pAN7-1 integration into chromosome Va. A replacement construct was therefore prepared, which contained sequences that flank the site of integration (Section 3.6.1). Integration of this construct into wild-type by a double crossover event would delete approximately 7.6 kb of genomic sequence known to be missing in the YI-20 mutant.

Analysis of the recombination outcomes following transformation of wild-type *P. paxilli*, with pPN1418 (Section 3.6.2 and 3.6.3) showed (i) 6.5% of transformants were single crossovers into the *Hin*dIII fragment, of which one transformant contained a deletion; (ii) 8.1% of transformants were single crossovers into the *Sst*I fragment, of which one transformant contained a deletion; (iii) 1.6% of transformants were double crossovers into both the *Hin*dIII and *Sst*I fragments, deleting the 7.6 kb between these fragments; and (iv) the remaining transformants had been integrated ectopically. There was uncertainty that a double crossover event would occur with this construct because of the size of the forced deletion. However, as 1.6% of transformants were the result of a double crossover event, this was obviously not a problem. Previous work in *P. paxilli* also showed that deletions were found at the site of homologous integration as a single crossover event at the brown spore locus (Itoh and Scott, 1994).

Analysis for paxilline, using the cELISA, showed that only two of the transformants containing homologous recombinations were Pax$^-$ and that these two mutants, CY-2 and CY-102, had deletions associated with a single crossover event. The two mutants that are the result of a double crossover, CY-35 and CY-39, produced as much paxilline as wild-type, confirming that this region was not involved with paxilline biosynthesis. As seen with YI-20, HPLC and TLC analysis showed the absence of paxilline and other expected intermediates for mutants CY-2 and CY-102, supporting an early block of the paxilline biosynthetic pathway. The presence of the ergosterol peak at 3.2 min suggests that the block in paxilline biosynthesis cannot be earlier than the formation of farnesyl pyrophosphate (FPP, Fig. 1.2) as ergosterol is synthesised from squalene, a 30 carbon terpenoid, which is initially formed by condensation of two FPP molecules. The literature also suggests (Barabato *et al.*, 1996; and Yiang *et al.*, 1995) that knocking out the synthesis of geranylgeranyl pyrophosphate (GGPP, Fig. 1.2) affects the growth of the culture.

Southern analysis of CY-2, CY-102 and CY-35 showed that, the integrating plasmid was at a single location but present as a tandem repeat (Section 3.6.3). This is likely due to transformation with a circular plasmid, as pPN1418 could not be linearised at a convenient site outside the two wild-type fragments. The deletions in CY-2, CY-102 and CY-35 were confirmed by Southern analysis. Further characterisation of CY-35 revealed a 7.6 kb deletion between the two wild-type fragments used in the pPN1418 construct. With only a small deletion and multiple copies of the integrating plasmid, the chromosome Va in CY-35 increases to a size similar to chromosome Vb (Section 3.8). The size range of the CY-2 and CY-102 deletions was estimated from chromosome separations (Section 3.8) and chromosome Va in these two mutants appears to be at least 80 kb smaller than the wild-type chromosome Va. However, as CY-2 and CY-102 contain multiple copies of the 10.8 kb plasmid, at least 54 kb (5 copies of pPN1418) of the deletion is replaced by plasmid. This result indicates that the size of the deletion in these two mutants are in the range of 120 - 150 kb, dependent on the number of integrating plasmids. However, the deletion could be as much as 100 kb larger, as sizing chromosomes has a large margin of error. It is clear from karyotype and Southern analysis of CY-2 and CY-102 that they contain deletions and do not have a chromosome rearrangement as seen in YI-20. As these two transformants are unable to produce paxilline, this gives strong evidence that the deleted sequences on chromosome Va are responsible for the Pax$^-$ phenotype observed. This also indicates that the region contained in the deletion is dispensable as there are no growth differences between wild-type and the mutants, CY-2 and CY-102. The large deletions found for both heterologous and homologous recombination are probably more common than the few reports to date would indicate (Kistler and Benny, 1992; Itoh and Scott, 1994). Deletions found as a result of heterologous plasmid integration in the plant

pathogenic fungus, *Nectria haematococca*, have been shown to be as large as 1000 kb and some strains appear to lack the entire chromosome into which the integration had occurred. These results indicate that the *Nectria haematococca* genome is extremely malleable and large portions may be non-essential for growth in culture (Kistler and Benny, 1992). It is possible that this locus, and the *brs* locus (Itoh and Scott, 1994) targeted in *P. paxilli* are both susceptible to deletion as appears to be the case for the penicillin biosynthetic cluster in *P. chrysogenum* (Fierro *et al.*, 1996)

Southern analysis of CY-2 revealed, unlike YI-20, the right-hand flanking sequence. Analysis of this right-hand flanking sequence in CY-2 with five restriction enzymes, shows all five enzyme sites map to a local distance of 100 bp. Although this could be possible, it could also indicate the deletion has gone as far as the telomere of the chromosome. However, this is difficult to prove and it appears from the data available that it should be possible to plasmid rescue such a sequence (Section 3.7 and 3.7.1). Rescue of this sequence would then enable a complete analysis of the size of the deletion in CY-2 and facilitate chromosome walking across the deletion from both directions. However, in every attempt to rescue right-hand flanking sequence, either by plasmid rescue or inverse PCR (IPCR), the only outcome was the rescue of the tandem repeat (Section 3.7 and 3.7.1). It is possible that the plasmid rescue of right-hand flanking sequence did not work due to insufficient promoter sequence adjacent to the ampicillin resistance gene. The IPCR, with a conservative placement of the pUCap primer in the β-lactamase gene, also gave a result expected from the tandem repeat and not right-hand flanking sequence. The difficulty in obtaining the single copy right-hand flanking sequence with the IPCR was thought to be due to the tandem repeat sequestering all the available primers, resulting only in amplification of a fragment from the tandem repeat and not of the required flanking sequence. However, if the deletion does reach the telomere of the chromosome, this would inhibit rescue of the flanking sequence by either of the above techniques.

In this study YI-20, a Pax$^-$ mutant of *P. paxilli*, created by plasmid tagging, is shown to be very complex, as it contains a large deletion and a chromosome rearrangement. It was important to establish that the tagged region of YI-20 was associated with paxilline biosynthesis. Homologous recombination was used to create two new Pax$^-$ mutants at the locus associated with pAN7-1 integration in YI-20. Although all three Pax$^-$ mutants have extensive deletions, the creation of the two mutants by homologous recombination indicates that the genes responsible for the phenotype are contained in the deletion.

It is now known from chromosome walking across the deletion that a gene cluster involved in paxilline biosynthesis is contained within the deletion seen in YI-20, CY-2 and CY-102 (Young and Scott, unpublished data). A series of deletion mutants created not only by the

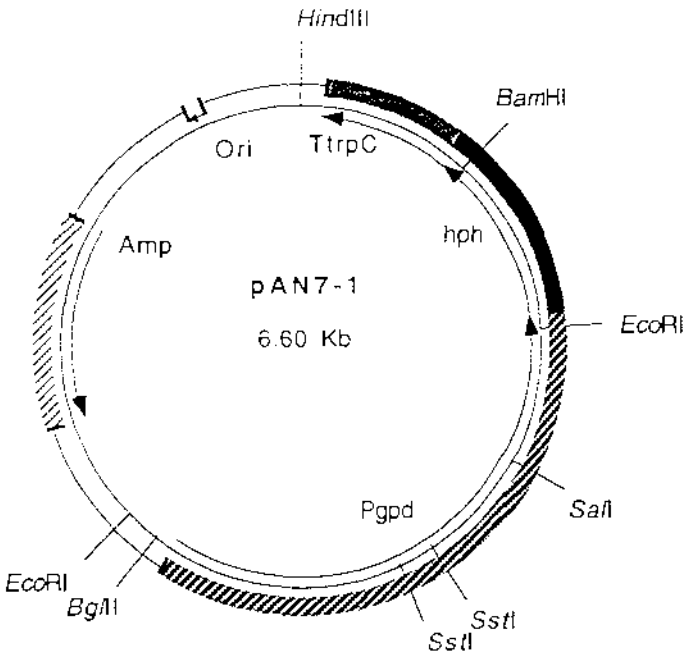
original illegitimate recombination of pAN7-1 but also by homologous recombination and restriction enzyme mediated integration (REMI), has narrowed the region of interest to approximately 70 kb (Appendix A3.1). Since this result, I have chromosome-walked 80 kb across the deletion (Appendix A3.1; Young and Scott, unpublished data). Spot sequencing of fragments contained in this deleted region showed sequence similarities to (i) geranylgeranyl pyrophosphate synthase (*pax*G), (ii) a P450 monooxygenase (*pax*P), (iii) a FAD dependent monooxygenase (*paxM*), (iv) a metabolite transport protein (*pax*T) and (v) a regulatory protein (*pax*R). All five genes are deleted in the three Pax$^-$ mutants, YI-20, CY-2 and CY-102. Finding geranylgeranyl pyrophosphate synthase associated within the deletion and proposed gene cluster was surprising. The literature indicates (Barabato *et al.*, 1996; and Yiang *et al.*, 1995) that this enzyme is required in primary metabolism, and that mutants lacking this gene are either not viable (Barabato *et al.*, 1996) or have impaired growth (Yiang *et al.*, 1995). A *pax*G knock out in *P. paxilli* using homologous recombination results in the expected Pax$^-$ phenotype and has no effect on the growth of the culture (McMillan, Young and Scott, unpublished data). Sequencing will continue across this 70 kb region to find other genes that may be involved in paxilline biosynthesis. RT-PCR is being used to confirm the expression pattern of each of these genes (Telfer and Scott, unpublished data) with a series of knock out studies in progress to determine the role of each gene (Young, McMillan and Scott, unpublished data).

Having now identified the paxilline biosynthetic gene cluster, future work will focus on (i) elucidation of the pathway, (ii) understanding the regulation of paxilline biosynthesis and (iii) identifying the function of each gene product. The complete sequence of this cluster, as well as Northern and RT-PCR analysis will identify other genes involved in the pathway as well as demarcate the left and right boundaries of the gene cluster. These boundaries will be defined as regions of the genome adjacent to the cluster, containing transcripts that are expressed under paxilline-inducing and non-paxilline-inducing conditions. Various molecular approaches can now be used to isolate the genes of closely related pathways, such as that of lolitrem B biosynthesis from the agriculturally important endophyte *Neotyphodium lolii*.

# *Appendix*
# *A1.0 Plasmid Maps*

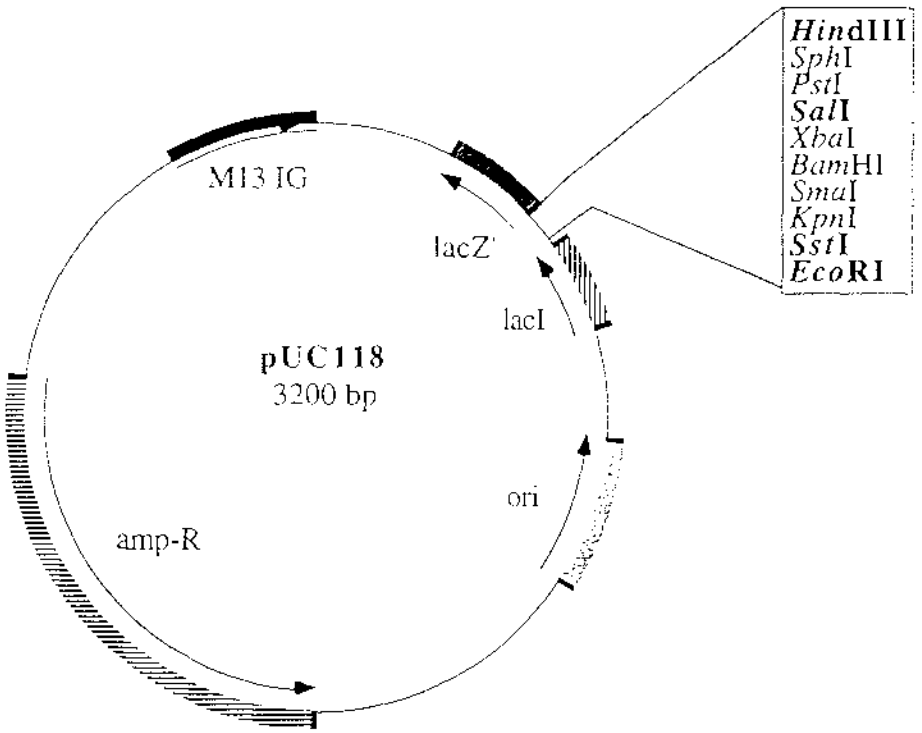
## *Appendix A1.1 pAN7-1*

Restriction map of pAN7-1.



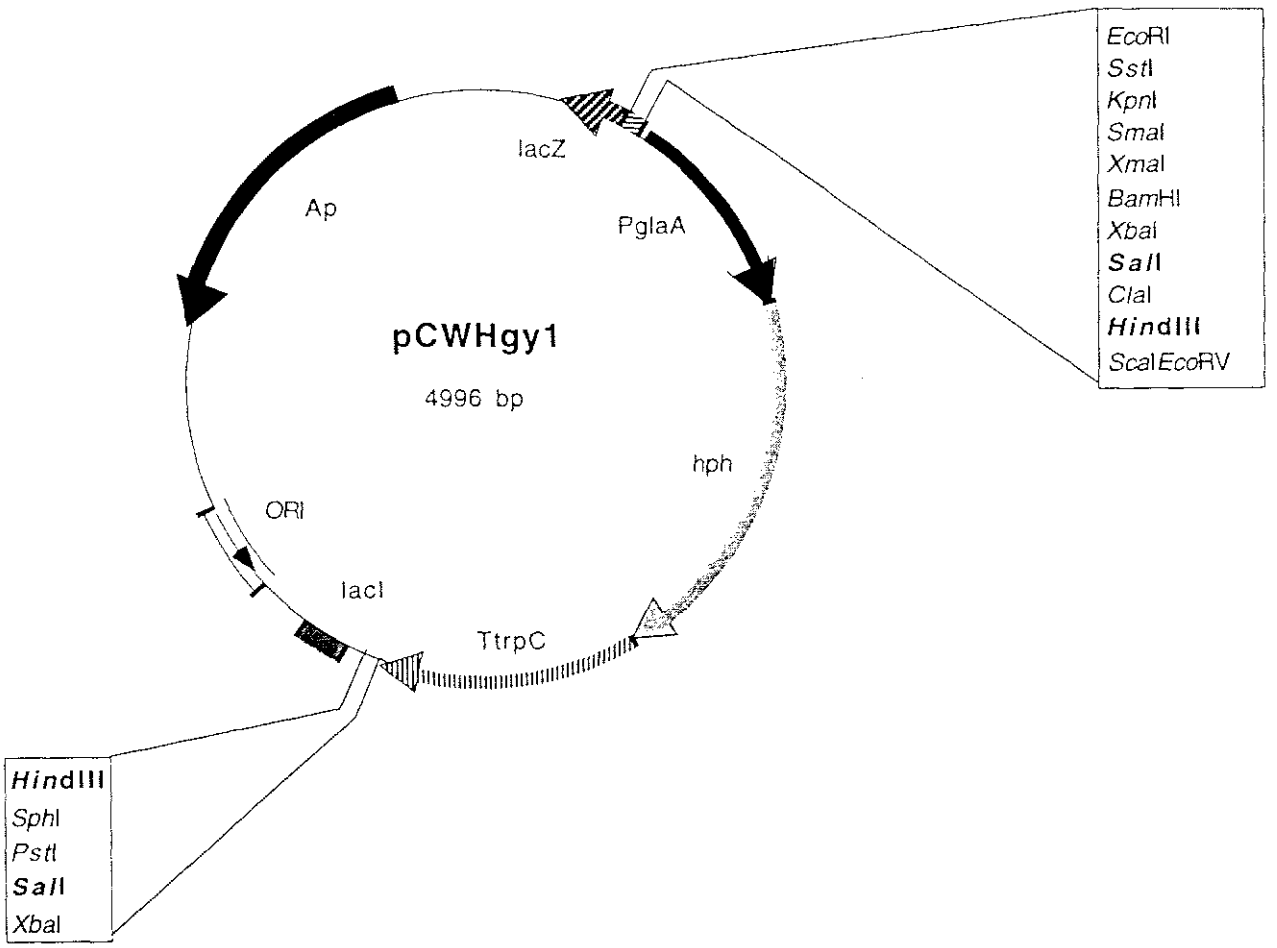
## *Appendix A1.2 pUC118*

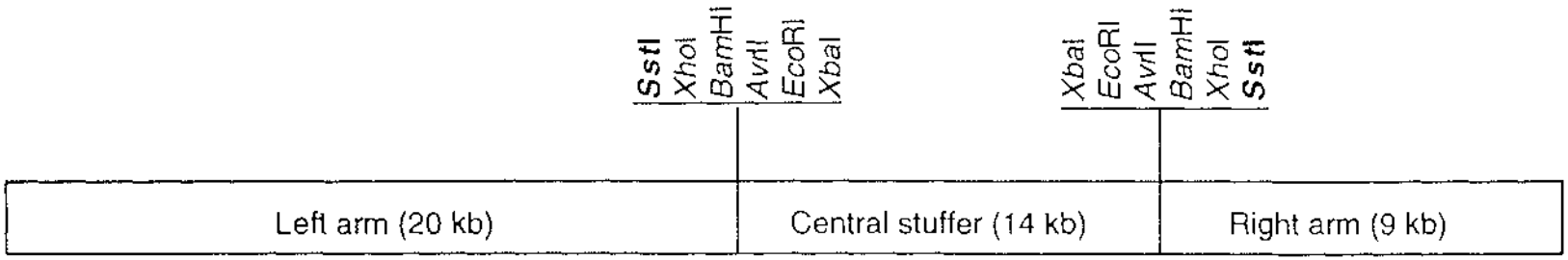
Restriction map of pUC118.



## *Appendix A1.3 pCWHyg1*

Restriction map of pCWHyg1.



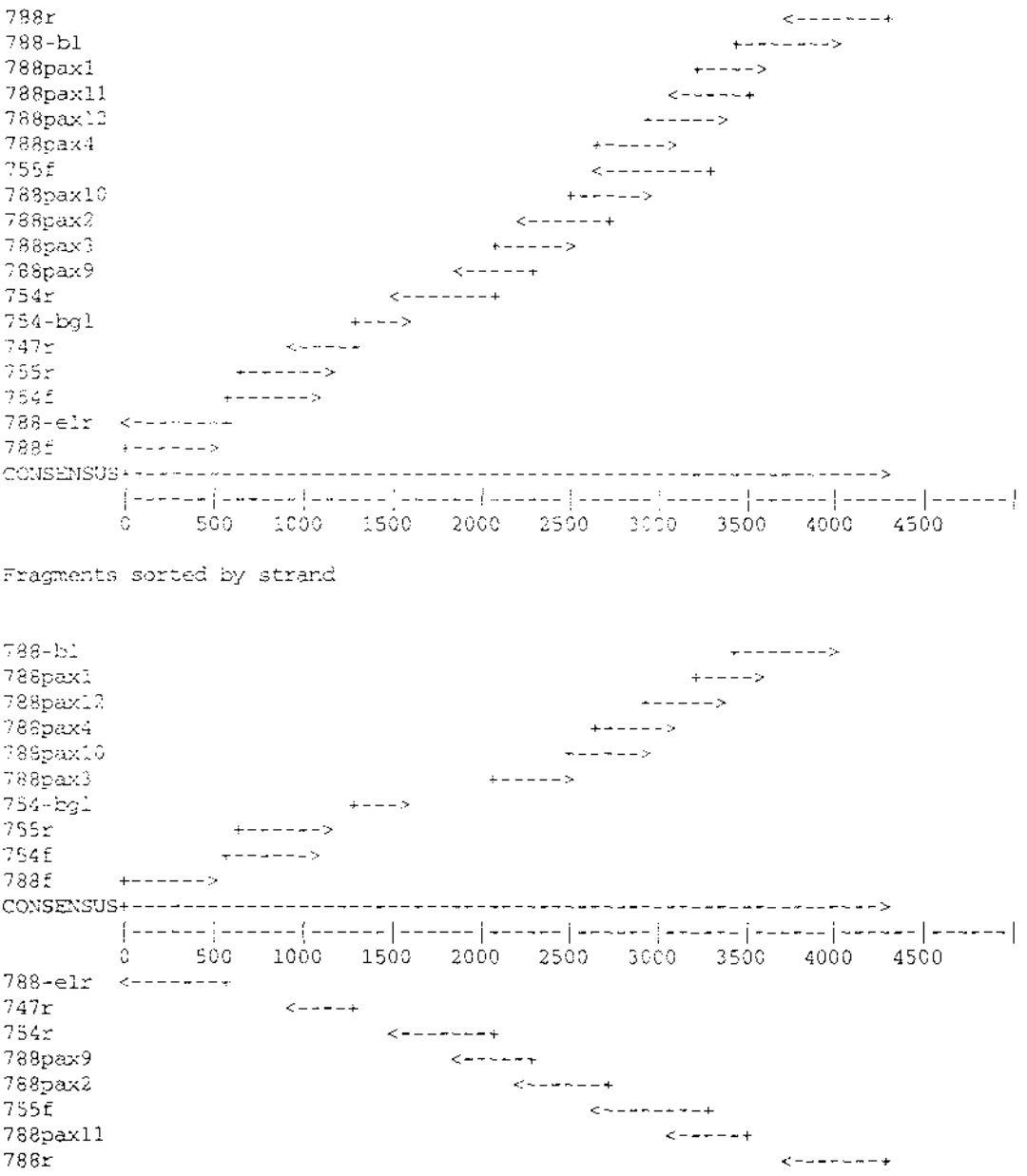



## *Appendix A1.4* λGEM-11

Restriction map showing SstI sites within the multiple cloning region of λGEM-11.

# *Appendix*
# *A2.0 Sequence Data*

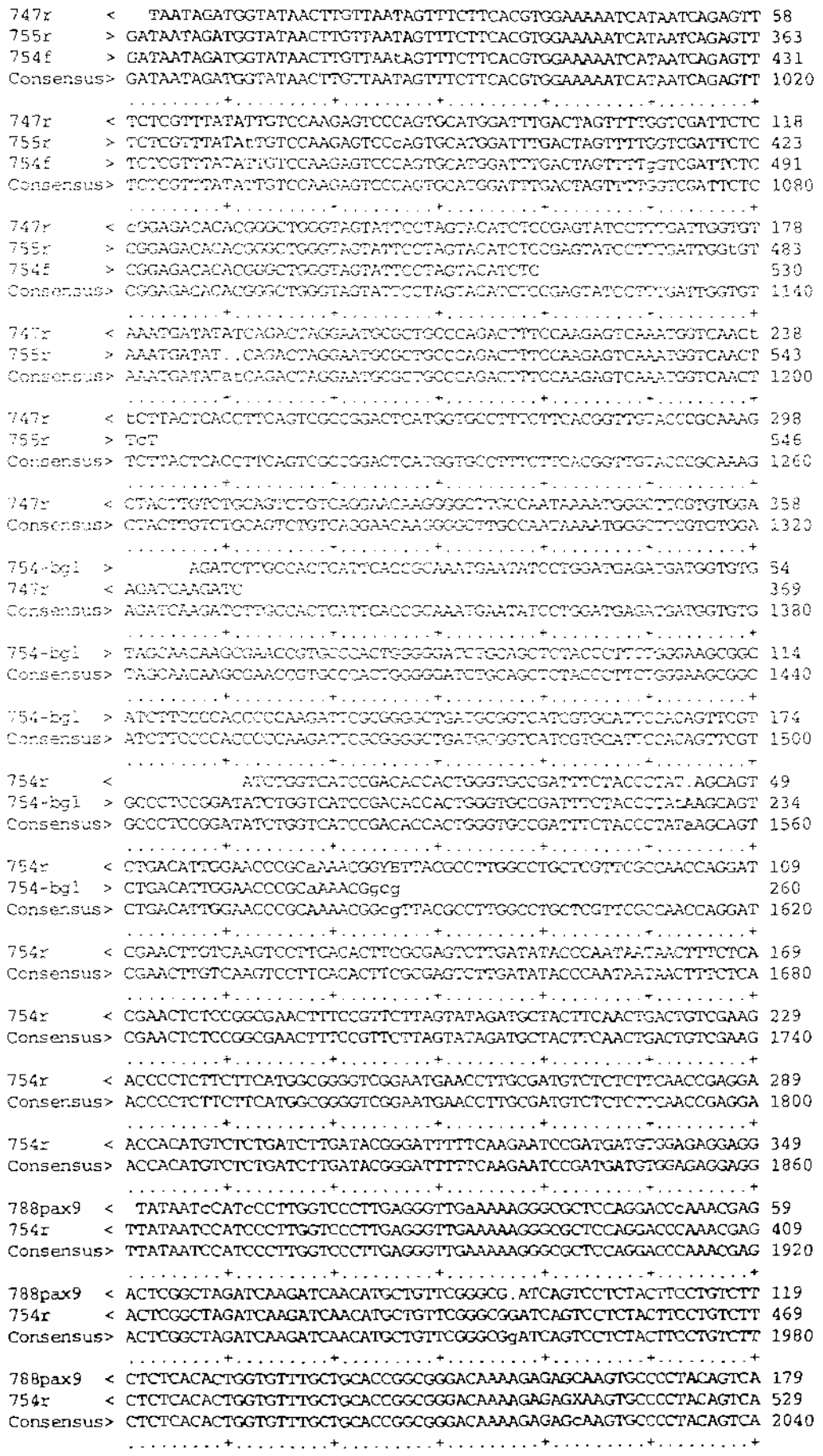
## *Appendix A2.1* *BIGPIC from pPN1408 (4.5 kb SstI fragment).*

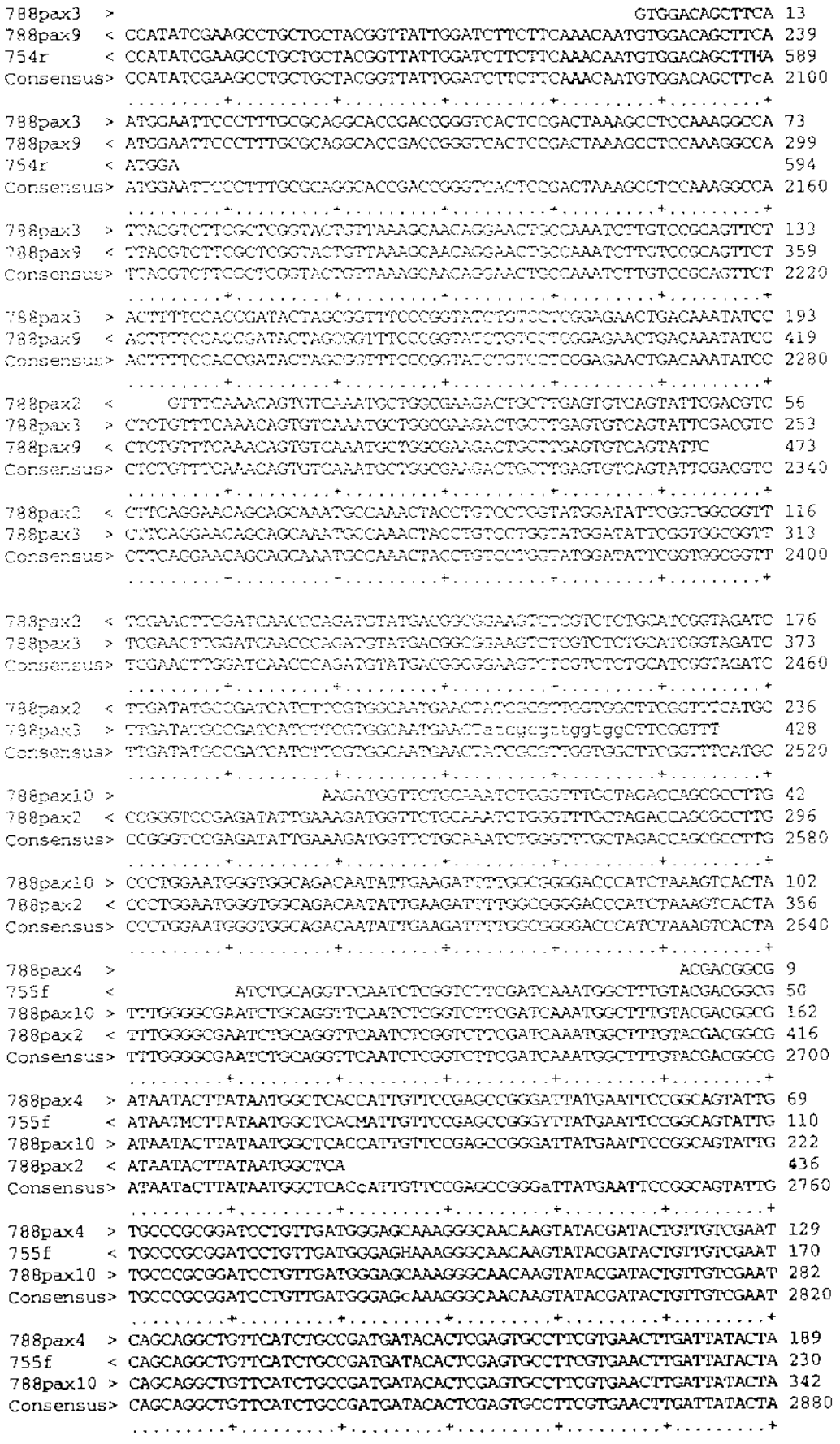


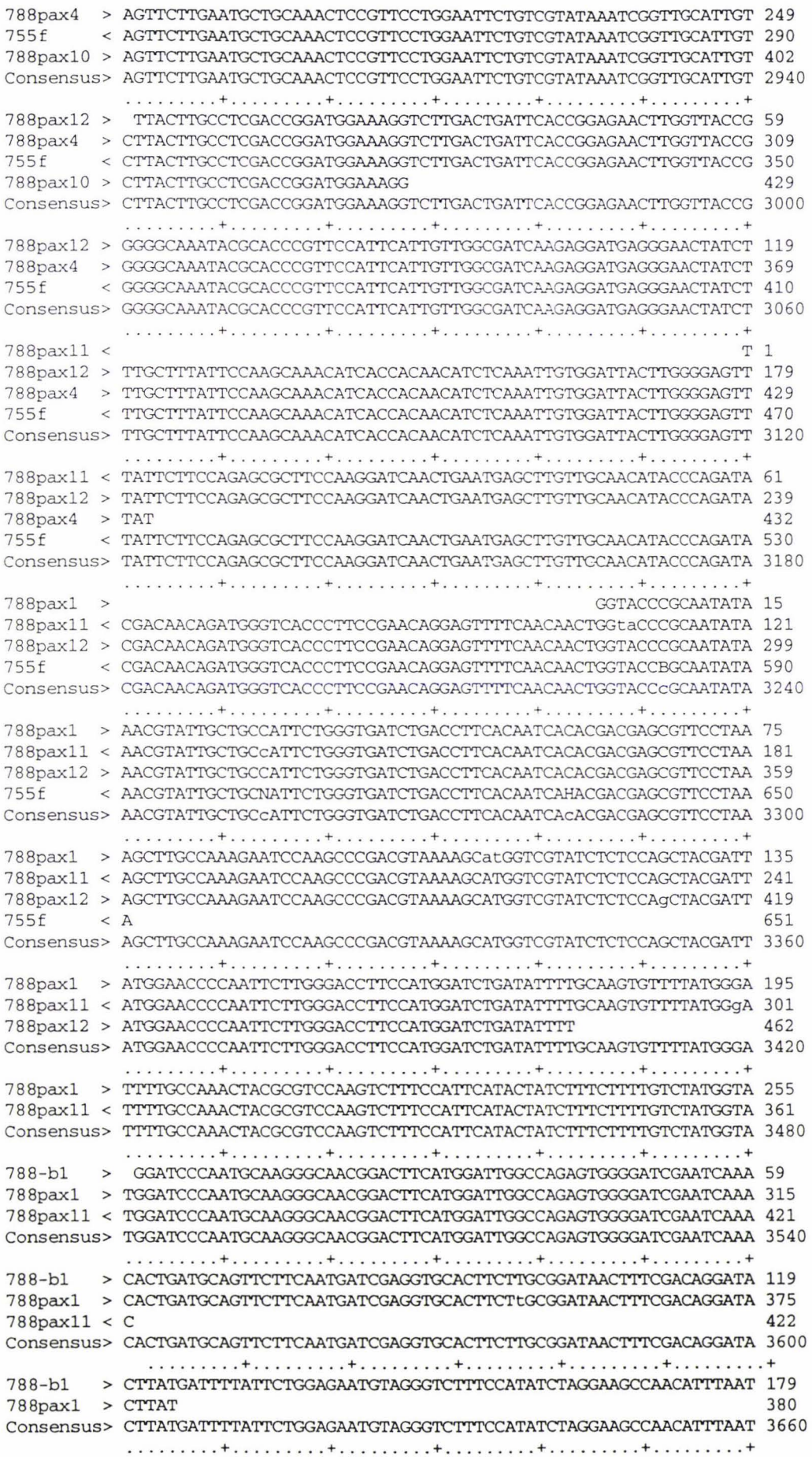
## *Appendix A2.2* *PRETTYOUT from pPN1408 contig.*

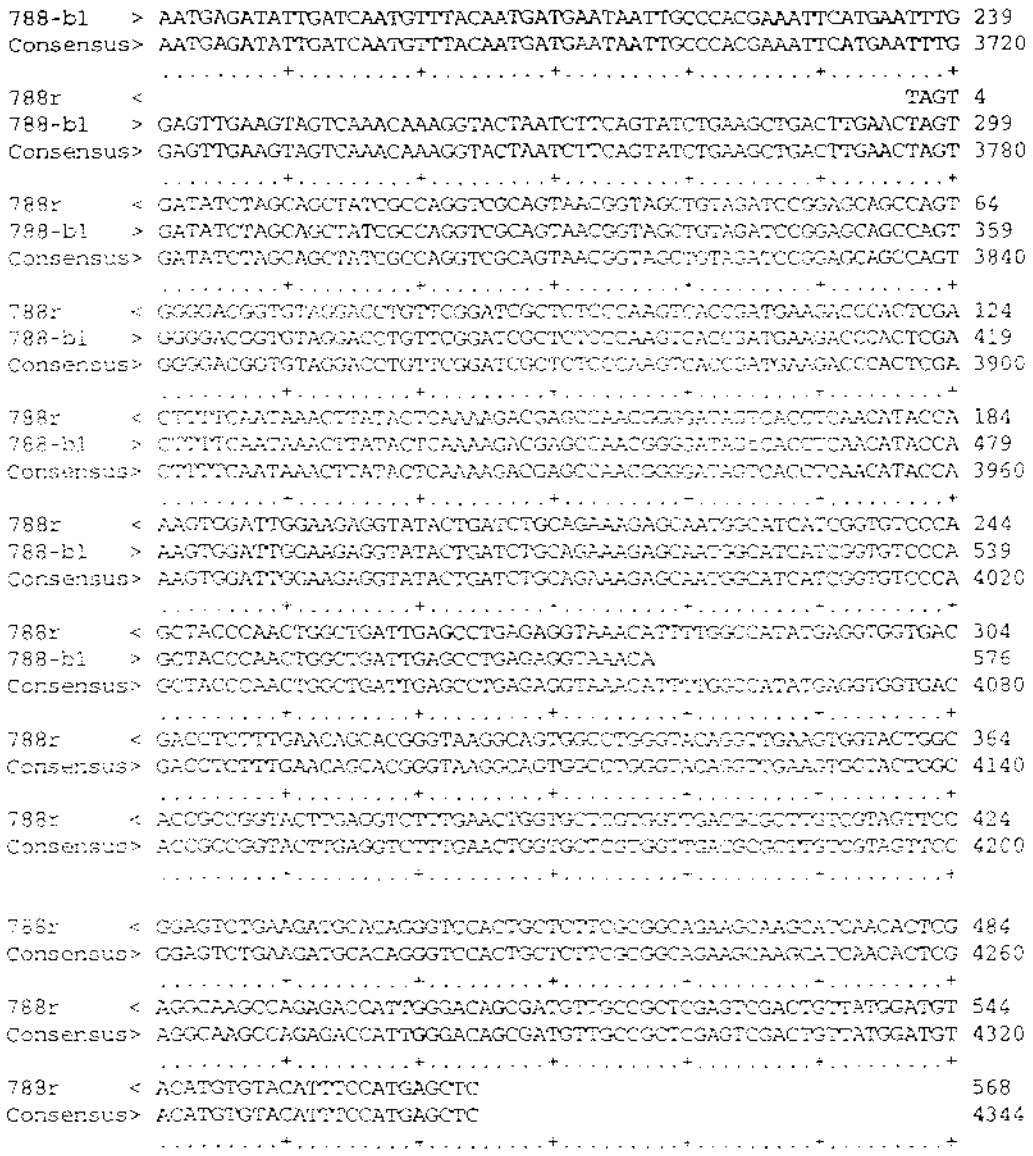
754 DNA is from pPN1385, 755 DNA is from pPN1387 and 788 DNA is from pPN1408.

```
788-elr  <                                                          GCTTGAAG 8
788f     > GAGCTCGCCGAGTTTCTGTTTGCCACGAATGGATATATAACATCCAGTCTTTGCTTGAAG 60
Consensus> GAGCTCGCCGAGTTTCTGTTTGCCACGAATGGATATATAACATCCAGTCTTTGCTTGAAG 60
           .........+.........+.........+.........+.........+.........+
788-elr  < CTGCAGTAATCTTGCAATCATCGATCAAAAGCATCACTGCACCTTGTCTTGCATCCACCA 68
788f     > TTGCAGTAATCTTGCAATCATCGATCAAAAGCATCACTGCACCTTGTCTTGCATCCACCA 120
Consensus> TTGCAGTAATCTTGCAATCATCGATCAAAAGCATCACTGCACCTTGTCTTGCATCCACCA 120
           .........+.........+.........+.........+.........+.........+
788-elr  < CTGGCAGTACTCCACATCTATCACACTACTACTGGATCAGCAACTCTACAAACCTGAGAA 128
788f     > CTGGCAGTACTCCACATCTATCACACTACTACTGGATCAGCAACTCTACAAACCTGAGAA 180
Consensus> CTGGCAGTACTCCACATCTATCACACTACTACTGGATCAGCAACTCTACAAACCTGAGAA 180
           .........+.........+.........+.........+.........+.........+
788-elr  < AGACGACTTACCTCCGCCATGCTGTCTTTACTCTCAGGAGCACAAACGCCCCAATCCGGA 188
788f     > AGACGACTTACCTCCGCCATGCTGTCTTTACTCTCAGGAGCACAAACGCCCCAATCCGGA 240
Consensus> AGACGACTTACCTCCGCCATGCTGTCTTTACTCTCAGGAGCACAAACGCCCCAATCCGGA 240
           .........+.........+.........+.........+.........+.........+
788-elr  < CGCACATCGCCTGGCAGTCACCACGATCATGAAACCTCCCACAACGAAACCCGCCATGTC 248
788f     > CGCACATCGCCTGGCAGTCACCACGATCATGAAACCTCCCACAACGAAACCCGCCATGTC 300
Consensus> CGCACATCGCCTGGCAGTCACCACGATCATGAAACCTCCCACAACGAAACCCGCCATGTC 300
           .........+.........+.........+.........+.........+.........+
788-elr  < TCAAGTCCACCCCGCAAGCGCATTGTGGTTGCCATGACCGGTGCCACCGGCGCAATGCTT 308
788f     > TCAAGTCCACCCCGCAAGCGCATTGTGGTTGCCATGACCGGTGCCACCGGCGCAATGCTT 360
Consensus> TCAAGTCCACCCCGCAAGCGCATTGTGGTTGCCATGACCGGTGCCACCGGCGCAATGCTT 360
           .........+.........+.........+.........+.........+.........+
788-elr  < GGTATCAAAGTCCTTATCGCCCTCCGCCGCCTCAACATCGAAACCCACCTGATCATGAGC 368
788f     > GGTATCAAAGTCCTTATCGCCCTCCGCCGCCTCAACATCGAAACCCACCTGATCATGAGC 420
Consensus> GGTATCAAAGTCCTTATCGCCCTCCGCCGCCTCAACATCGAAACCCACCTGATCATGAGC 420
           .........+.........+.........+.........+.........+.........+
788-elr  < AAATGGGCCGAAGCAACAATTAAATACGAGACCGATTATCACCCATCAAACGTCAAGGCC 428
788f     > AAATGGGCCGAAGCAACAATTAAATACGAGACCGATTATCACCCATCAAACGTCAAGGCC 480
Consensus> AAATGGGCCGAAGCAACAATTAAATACGAGACCGATTATCACCCATCAAACGTCAAGGCC 480
           .........+.........+.........+.........+.........+.........+
788-elr  < CTTGCTGATCACGTCCACAACATCAACGATATGGCGGCGCCGGTATCGAGTGGTTCATTC 488
788f     > CTTGCTGATCACGTCCACAAC                                        501
Consensus> CTTGCTGATCACGTCCACAACATCAACGATATGGCGGCGCCGGTATCGAGTGGTTCATTC 540
           .........+.........+.........+.........+.........+.........+
754f     >                                                   GAATTCATAGT 11
788-elr  < AAGACAGACGGGATGATTATTGTTCCTTGCAGCATGAAGACACTTGCTGGAATTC       543
Consensus> AAGACAGACGGGATGATTATTGTTCCTTGCAGCATGAAGACACTTGCTGGAATTCATAGT 600
           .........+.........+.........+.........+.........+.........+
755r     >                                                         AAG 3
754f     > GGATTCTGCGATGACTTGATTTCGCGGACAGCAGATGTCATGCTCAAAGAGCGACGCAAG 71
Consensus> GGATTCTGCGATGACTTGATTTCGCGGACAGCAGATGTCATGCTCAAAGAGCGACGCAAG 660
           .........+.........+.........+.........+.........+.........+
755r     > CTTGTCCTTGTTGCGAGAGAGACTCCGCTGAGTGATATTCACCTGCGGAATATGCTTGAG 63
754f     > CTTGTCCTTGTTGCGAGAGAGACTCCGCTGAGTGATATTCACCTGCGGAATATGCTTGAG 131
Consensus> CTTGTCCTTGTTGCGAGAGAGACTCCGCTGAGTGATATTCACCTGCGGAATATGCTTGAG 720
           .........+.........+.........+.........+.........+.........+
755r     > ATCACTCGAAGTGGCGCCATTATTTTCCCACCTGTTCCGGCGTACTATATTCGAGCTGCC 123
754f     > ATCACTCGAAGTGGCGCCATTATTTTCCCACCTGTTCCGGCGTACTATATTCGAGCTGCC 191
Consensus> ATCACTCGAAGTGGCGCCATTATTTTCCCACCTGTTCCGGCGTACTATATTCGAGCTGCC 780
           .........+.........+.........+.........+.........+.........+
755r     > TCGGTTGATGACTTGATCAATCAAAGTGTTGGACGAATGCTAGATCTGTTTGATTTGGAC 183
754f     > TCGGTTGATGACTTGATCAATCAAAGTGTTGGACGAATGCTAGATCTGTTTGATTTGGAC 251
Consensus> TCGGTTGATGACTTGATCAATCAAAGTGTTGGACGAATGCTAGATCTGTTTGATTTGGAC 840
           .........+.........+.........+.........+.........+.........+
755r     > ACGGGCGATTTCGAGCGTTGGGATGGGTGGCAGACTGAGAAATGAGCTACCTAGAGTGCA 243
754f     > ACGGGCGATTTCGAGCGTTGGGATGGGTGGCAGACTGAGAAATGAGCTACCTAGAGTGCA 311
Consensus> ACGGGCGATTTCGAGCGTTGGGATGGGTGGCAGACTGAGAAATGAGCTACCTAGAGTGCA 900
           .........+.........+.........+.........+.........+.........+
755r     > CGGGGAAACGGAGTTTTCCGGAAATGAACTTTTTTGCATTGCATATTTAGCGTTATGAGG 303
754f     > CGGGGAAACGGAGTTTTCCgGAAATGAACTTTTTTGCATTGCATATTTAGCGTTATGAGG 371
Consensus> CGGGGAAACGGAGTTTTCCGGAAATGAACTTTTTTGCATTGCATATTTAGCGTTATGAGG 960
           .........+.........+.........+.........+.........+.........+
```


747r < TAATAGATGGTATAACTTGTTAATAGTTTCTTCACGTGGAAAAATCATAATCAGAGTT 58
755r > GATAATAGATGGTATAACTTGTTAATAGTTTCTTCACGTGGAAAAATCATAATCAGAGTT 363
754f > GATAATAGATGGTATAACTTGTTAATAGTTTCTTCACGTGGAAAAATCATAATCAGAGTT 431
Consensus> GATAATAGATGGTATAACTTGTTAATAGTTTCTTCACGTGGAAAAATCATAATCAGAGTT 1020
747r < TCTCGTTTATATTGTCCAAGAGTCCCAGTGCATGGATTTGACTAGTTTTGGTCGATTCTC 118
755r > TCTCGTTTATAttGTCCAAGAGTCCcAGTGCATGGATTTGACTAGTTTTGGTCGATTCTC 423
754f > TCTCGTTTATATTGTCCAAGAGTCCCAGTGCATGGATTTGACTAGTTTTgGTCGATTCTC 491
Consensus> TCTCGTTTATATTGTCCAAGAGTCCCAGTGCATGGATTTGACTAGTTTTGGTCGATTCTC 1080
747r < cGGAGACACACGGGCTGGGTAGTATTCCTAGTACATCTCCGAGTATCCTTTGATTGGTGT 178
755r > CGGAGACACACGGGCTGGGTAGTATTCCTAGTACATCTCCGAGTATCCTTTGATTGGtGT 483
754f > CGGAGACACACGGGCTGGGTAGTATTCCTAGTACATCTC 530
Consensus> CGGAGACACACGGGCTGGGTAGTATTCCTAGTACATCTCCGAGTATCCTTTGATTGGTGT 1140
747r < AAATGATATATCAGACTAGGAATGCGCTGCCCAGACTTTCCAAGAGTCAAATGGTCAACt 238
755r > AAATGATAT..CAGACTAGGAATGCGCTGCCCAGACTTTCCAAGAGTCAAATGGTCAACT 543
Consensus> AAATGATATatCAGACTAGGAATGCGCTGCCCAGACTTTCCAAGAGTCAAATGGTCAACT 1200
747r < tCTTACTCACCTTCAGTCGCCGGACTCATGGTGCCTTTCTTCACGGTTGTACCCGCAAAG 298
755r > TcT 546
Consensus> TCTTACTCACCTTCAGTCGCCGGACTCATGGTGCCTTTCTTCACGGTTGTACCCGCAAAG 1260
747r < CTACTTGTCTGCAGTCTGTCAGGAACAAGGGGCTTGCCAATAAAATGGGCTTCGTGTGGA 358
Consensus> CTACTTGTCTGCAGTCTGTCAGGAACAAGGGGCTTGCCAATAAAATGGGCTTCGTGTGGA 1320
754-bgl > AGATCTTGCCACTCATTCACCGCAAATGAATATCCTGGATGAGATGATGGTGTG 54
747r < AGATCAAGATC 369
Consensus> AGATCAAGATCTTGCCACTCATTCACCGCAAATGAATATCCTGGATGAGATGATGGTGTG 1380
754-bgl > TAGCAACAAGCGAACCGTGCCCACTGGGGGATCTGCAGCTCTACCCTTCTGGGAAGCGGC 114
Consensus> TAGCAACAAGCGAACCGTGCCCACTGGGGGATCTGCAGCTCTACCCTTCTGGGAAGCGGC 1440
754-bgl > ATCTTCCCCACCCCCAAGATTCGCGGGGCTGATGCGGTCATCGTGCATTCCACAGTTCGT 174
Consensus> ATCTTCCCCACCCCCAAGATTCGCGGGGCTGATGCGGTCATCGTGCATTCCACAGTTCGT 1500
754r < ATCTGGTCATCCGACACCACTGGGTGCCGATTTCTACCCTAT.AGCAGT 49
754-bgl > GCCCTCCGGATATCTGGTCATCCGACACCACTGGGTGCCGATTTCTACCCTATaAGCAGT 234
Consensus> GCCCTCCGGATATCTGGTCATCCGACACCACTGGGTGCCGATTTCTACCCTATaAGCAGT 1560
754r < CTGACATTGGAACCCGCaAAACGGYBTTACGCCTTGGCCTGCTCGTTCGCCAACCAGGAT 109
754-bgl > CTGACATTGGAACCCGCaAAACGGcg 260
Consensus> CTGACATTGGAACCCGCAAAACGGcgTTACGCCTTGGCCTGCTCGTTCGCCAACCAGGAT 1620
754r < CGAACTTGTCAAGTCCTTCACACTTCGCGAGTCTTGATATACCCAATAATAACTTTCTCA 169
Consensus> CGAACTTGTCAAGTCCTTCACACTTCGCGAGTCTTGATATACCCAATAATAACTTTCTCA 1680
754r < CGAACTCTCCGGCGAACTTTCCGTTCTTAGTATAGATGCTACTTCAACTGACTGTCGAAG 229
Consensus> CGAACTCTCCGGCGAACTTTCCGTTCTTAGTATAGATGCTACTTCAACTGACTGTCGAAG 1740
754r < ACCCCTCTTCTTCATGGCGGGGTCGGAATGAACCTTGCGATGTCTCTCTTCAACCGAGGA 289
Consensus> ACCCCTCTTCTTCATGGCGGGGTCGGAATGAACCTTGCGATGTCTCTCTTCAACCGAGGA 1800
754r < ACCACATGTCTCTGATCTTGATACGGGATTTTTCAAGAATCCGATGATGTGGAGAGGAGG 349
Consensus> ACCACATGTCTCTGATCTTGATACGGGATTTTTCAAGAATCCGATGATGTGGAGAGGAGG 1860
788pax9 < TATAATcCATcCCTTGGTCCCTTGAGGGTTGaAAAAGGGCGCTCCAGGACCcAAACGAG 59
754r < TTATAATCCATCCCTTGGTCCCTTGAGGGTTGAAAAAGGGCGCTCCAGGACCCAAACGAG 409
Consensus> TTATAATCCATCCCTTGGTCCCTTGAGGGTTGAAAAAGGGCGCTCCAGGACCCAAACGAG 1920
788pax9 < ACTCGGCTAGATCAAGATCAACATGCTGTTCGGGCG.ATCAGTCCTCTACTTCCTGTCTT 119
754r < ACTCGGCTAGATCAAGATCAACATGCTGTTCGGGCGGATCAGTCCTCTACTTCCTGTCTT 469
Consensus> ACTCGGCTAGATCAAGATCAACATGCTGTTCGGGCGgATCAGTCCTCTACTTCCTGTCTT 1980
788pax9 < CTCTCACACTGGTGTTTGCTGCACCGGCGGGACAAAAGAGAGCAAGTGCCCCTACAGTCA 179
754r < CTCTCACACTGGTGTTTGCTGCACCGGCGGGACAAAAGAGAGXAAGTGCCCCTACAGTCA 529
Consensus> CTCTCACACTGGTGTTTGCTGCACCGGCGGGACAAAAGAGAGcAAGTGCCCCTACAGTCA 2040


788pax3 > GTGGACAGCTTCA 13
788pax9 < CCATATCGAAGCCTGCTGCTACGGTTATTGGATCTTCTTCAAACAATGTGGACAGCTTCA 239
754r < CCATATCGAAGCCTGCTGCTACGGTTATTGGATCTTCTTCAAACAATGTGGACAGCTTHA 589
Consensus> CCATATCGAAGCCTGCTGCTACGGTTATTGGATCTTCTTCAAACAATGTGGACAGCTTcA 2100
788pax3 > ATGGAATTCCCTTTGCGCAGGCACCGACCGGGTCACTCCGACTAAAGCCTCCAAAGGCCA 73
788pax9 < ATGGAATTCCCTTTGCGCAGGCACCGACCGGGTCACTCCGACTAAAGCCTCCAAAGGCCA 299
754r < ATGGA 594
Consensus> ATGGAATTCCCTTTGCGCAGGCACCGACCGGGTCACTCCGACTAAAGCCTCCAAAGGCCA 2160
788pax3 > TTACGTCTTCGCTCGGTACTGTTAAAGCAACAGGAACTGCCAAATCTTGTCCGCAGTTCT 133
788pax9 < TTACGTCTTCGCTCGGTACTGTTAAAGCAACAGGAACTGCCAAATCTTGTCCGCAGTTCT 359
Consensus> TTACGTCTTCGCTCGGTACTGTTAAAGCAACAGGAACTGCCAAATCTTGTCCGCAGTTCT 2220
788pax3 > ACTTTTCCACCGATACTAGCGGTTTCCCGGTATCTGTCCTCGGAGAACTGACAAATATCC 193
788pax9 < ACTTTTCCACCGATACTAGCGGTTTCCCGGTATCTGTCCTCGGAGAACTGACAAATATCC 419
Consensus> ACTTTTCCACCGATACTAGCGGTTTCCCGGTATCTGTCCTCGGAGAACTGACAAATATCC 2280
788pax2 < GTTTCAAACAGTGTCAAATGCTGGCGAAGACTGCTTGAGTGTCAGTATTCGACGTC 56
788pax3 > CTCTGTTTCAAACAGTGTCAAATGCTGGCGAAGACTGCTTGAGTGTCAGTATTCGACGTC 253
788pax9 < CTCTGTTTCAAACAGTGTCAAATGCTGGCGAAGACTGCTTGAGTGTCAGTATTC 473
Consensus> CTCTGTTTCAAACAGTGTCAAATGCTGGCGAAGACTGCTTGAGTGTCAGTATTCGACGTC 2340
788pax3 < CTTCAGGAACAGCAGCAAATGCCAAACTACCTGTCCTGGTATGGATATTCGGTGGCGGTT 116
788pax3 > CTTCAGGAACAGCAGCAAATGCCAAACTACCTGTCCTGGTATGGATATTCGGTGGCGGTT 313
Consensus> CTTCAGGAACAGCAGCAAATGCCAAACTACCTGTCCTGGTATGGATATTCGGTGGCGGTT 2400
788pax2 < TCGAACTTCGATCAACCCAGATGTATGACGGCGGAAGTCTCGTCTCTGCATCGGTAGATC 176
788pax3 > TCGAACTTCGATCAACCCAGATGTATGACGGCGGAAGTCTCGTCTCTGCATCGGTAGATC 373
Consensus> TCGAACTTCGATCAACCCAGATGTATGACGGCGGAAGTCTCGTCTCTGCATCGGTAGATC 2460
788pax2 < TTGATATGCCGATCATCTTCGTGGCAATGAACTATCGCGTTGGTGGCTTCGGTTTCATGC 236
788pax3 > TTGATATGCCGATCATCTTCGTGGCAATGAACTatcgcgttggtggCTTCGGTTT 428
Consensus> TTGATATGCCGATCATCTTCGTGGCAATGAACTATCGCGTTGGTGGCTTCGGTTTCATGC 2520
788pax10 > AAGATGGTTCTGCAAATCTGGGTTTGCTAGACCAGCGCCTTG 42
788pax2 < CCGGGTCCGAGATATTGAAAGATGGTTCTGCAAATCTGGGTTTGCTAGACCAGCGCCTTG 296
Consensus> CCGGGTCCGAGATATTGAAAGATGGTTCTGCAAATCTGGGTTTGCTAGACCAGCGCCTTG 2580
788pax10 > CCCTGGAATGGGTGGCAGACAATATTGAAGATTTTGGCGGGGACCCATCTAAAGTCACTA 102
788pax2 < CCCTGGAATGGGTGGCAGACAATATTGAAGATTTTGGCGGGGACCCATCTAAAGTCACTA 356
Consensus> CCCTGGAATGGGTGGCAGACAATATTGAAGATTTTGGCGGGGACCCATCTAAAGTCACTA 2640
788pax4 > ACGACGGCG 9
755f < ATCTGCAGGTTCAATCTCGGTCTTCGATCAAATGGCTTTGTACGACGGCG 50
788pax10 > TTTGGGGCGAATCTGCAGGTTCAATCTCGGTCTTCGATCAAATGGCTTTGTACGACGGCG 162
788pax2 < TTTGGGGCGAATCTGCAGGTTCAATCTCGGTCTTCGATCAAATGGCTTTGTACGACGGCG 416
Consensus> TTTGGGGCGAATCTGCAGGTTCAATCTCGGTCTTCGATCAAATGGCTTTGTACGACGGCG 2700
788pax4 > ATAATACTTATAATGGCTCACCATTGTTCCGAGCCGGGATTATGAATTCCGGCAGTATTG 69
755f < ATAATMCTTATAATGGCTCACMATTGTTCCGAGCCGGGYTTATGAATTCCGGCAGTATTG 110
788pax10 > ATAATACTTATAATGGCTCACCATTGTTCCGAGCCGGGATTATGAATTCCGGCAGTATTG 222
788pax2 < ATAATACTTATAATGGCTCA 436
Consensus> ATAATaCTTATAATGGCTCACcATTGTTCCGAGCCGGGaTTATGAATTCCGGCAGTATTG 2760
788pax4 > TGCCCGCGGATCCTGTTGATGGGAGCAAAGGGCAACAAGTATACGATACTGTTGTCGAAT 129
755f < TGCCCGCGGATCCTGTTGATGGGAGHAAAGGGCAACAAGTATACGATACTGTTGTCGAAT 170
788pax10 > TGCCCGCGGATCCTGTTGATGGGAGCAAAGGGCAACAAGTATACGATACTGTTGTCGAAT 282
Consensus> TGCCCGCGGATCCTGTTGATGGGAGcAAAGGGCAACAAGTATACGATACTGTTGTCGAAT 2820
788pax4 > CAGCAGGCTGTTCATCTGCCGATGATACACTCGAGTGCCTTCGTGAACTTGATTATACTA 189
755f < CAGCAGGCTGTTCATCTGCCGATGATACACTCGAGTGCCTTCGTGAACTTGATTATACTA 230
788pax10 > CAGCAGGCTGTTCATCTGCCGATGATACACTCGAGTGCCTTCGTGAACTTGATTATACTA 342
Consensus> CAGCAGGCTGTTCATCTGCCGATGATACACTCGAGTGCCTTCGTGAACTTGATTATACTA 2880


788pax4 > AGTTCTTGAATGCTGCAAACTCCGTTCCTGGAATTCTGTCGTATAAATCGGTTGCATTGT 249
755f < AGTTCTTGAATGCTGCAAACTCCGTTCCTGGAATTCTGTCGTATAAATCGGTTGCATTGT 290
788pax10 > AGTTCTTGAATGCTGCAAACTCCGTTCCTGGAATTCTGTCGTATAAATCGGTTGCATTGT 402
Consensus> AGTTCTTGAATGCTGCAAACTCCGTTCCTGGAATTCTGTCGTATAAATCGGTTGCATTGT 2940
788pax12 > TTACTTGCCTCGACCGGATGGAAAGGTCTTGACTGATTCACCGGAGAACTTGGTTACCG 59
788pax4 > CTTACTTGCCTCGACCGGATGGAAAGGTCTTGACTGATTCACCGGAGAACTTGGTTACCG 309
755f < CTTACTTGCCTCGACCGGATGGAAAGGTCTTGACTGATTCACCGGAGAACTTGGTTACCG 350
788pax10 > CTTACTTGCCTCGACCGGATGGAAAGG 429
Consensus> CTTACTTGCCTCGACCGGATGGAAAGGTCTTGACTGATTCACCGGAGAACTTGGTTACCG 3000
788pax12 > GGGGCAAATACGCACCCGTTCCATTCATTGTTGGCGATCAAGAGGATGAGGGAACTATCT 119
788pax4 > GGGGCAAATACGCACCCGTTCCATTCATTGTTGGCGATCAAGAGGATGAGGGAACTATCT 369
755f < GGGGCAAATACGCACCCGTTCCATTCATTGTTGGCGATCAAGAGGATGAGGGAACTATCT 410
Consensus> GGGGCAAATACGCACCCGTTCCATTCATTGTTGGCGATCAAGAGGATGAGGGAACTATCT 3060
788pax11 < T 1
788pax12 > TTGCTTTATTCCAAGCAAACATCACCACAACATCTCAAATTGTGGATTACTTGGGGAGTT 179
788pax4 > TTGCTTTATTCCAAGCAAACATCACCACAACATCTCAAATTGTGGATTACTTGGGGAGTT 429
755f < TTGCTTTATTCCAAGCAAACATCACCACAACATCTCAAATTGTGGATTACTTGGGGAGTT 470
Consensus> TTGCTTTATTCCAAGCAAACATCACCACAACATCTCAAATTGTGGATTACTTGGGGAGTT 3120
788pax11 < TATTCTTCCAGAGCGCTTCCAAGGATCAACTGAATGAGCTTGTTGCAACATACCCAGATA 61
788pax12 > TATTCTTCCAGAGCGCTTCCAAGGATCAACTGAATGAGCTTGTTGCAACATACCCAGATA 239
788pax4 > TAT 432
755f < TATTCTTCCAGAGCGCTTCCAAGGATCAACTGAATGAGCTTGTTGCAACATACCCAGATA 530
Consensus> TATTCTTCCAGAGCGCTTCCAAGGATCAACTGAATGAGCTTGTTGCAACATACCCAGATA 3180
788pax1 > GGTACCCGCAATATA 15
788pax11 < CGACAACAGATGGGTCACCCTTCCGAACAGGAGTTTTCAACAACTGGtaCCCGCAATATA 121
788pax12 > CGACAACAGATGGGTCACCCTTCCGAACAGGAGTTTTCAACAACTGGTACCCGCAATATA 299
755f < CGACAACAGATGGGTCACCCTTCCGAACAGGAGTTTTCAACAACTGGTACCBGCAATATA 590
Consensus> CGACAACAGATGGGTCACCCTTCCGAACAGGAGTTTTCAACAACTGGTACCcGCAATATA 3240
788pax1 > AACGTATTGCTGCCATTCTGGGTGATCTGACCTTCACAATCACACGACGAGCGTTCCTAA 75
788pax11 < AACGTATTGCTGCcATTCTGGGTGATCTGACCTTCACAATCACACGACGAGCGTTCCTAA 181
788pax12 > AACGTATTGCTGCCATTCTGGGTGATCTGACCTTCACAATCACACGACGAGCGTTCCTAA 359
755f < AACGTATTGCTGCNATTCTGGGTGATCTGACCTTCACAATCAHACGACGAGCGTTCCTAA 650
Consensus> AACGTATTGCTGCcATTCTGGGTGATCTGACCTTCACAATCAcACGACGAGCGTTCCTAA 3300
788pax1 > AGCTTGCCAAAGAATCCAAGCCCGACGTAAAAGCatGGTCGTATCTCTCCAGCTACGATT 135
788pax11 < AGCTTGCCAAAGAATCCAAGCCCGACGTAAAAGCATGGTCGTATCTCTCCAGCTACGATT 241
788pax12 > AGCTTGCCAAAGAATCCAAGCCCGACGTAAAAGCATGGTCGTATCTCTCCAgCTACGATT 419
755f < A 651
Consensus> AGCTTGCCAAAGAATCCAAGCCCGACGTAAAAGCATGGTCGTATCTCTCCAGCTACGATT 3360
788pax1 > ATGGAACCCCAATTCTTGGGACCTTCCATGGATCTGATATTTTGCAAGTGTTTTATGGGA 195
788pax11 < ATGGAACCCCAATTCTTGGGACCTTCCATGGATCTGATATTTTGCAAGTGTTTTATGGgA 301
788pax12 > ATGGAACCCCAATTCTTGGGACCTTCCATGGATCTGATATTTT 462
Consensus> ATGGAACCCCAATTCTTGGGACCTTCCATGGATCTGATATTTTGCAAGTGTTTTATGGGA 3420
788pax1 > TTTTGCCAAACTACGCGTCCAAGTCTTTCCATTCATACTATCTTTCTTTTGTCTATGGTA 255
788pax11 < TTTTGCCAAACTACGCGTCCAAGTCTTTCCATTCATACTATCTTTCTTTTGTCTATGGTA 361
Consensus> TTTTGCCAAACTACGCGTCCAAGTCTTTCCATTCATACTATCTTTCTTTTGTCTATGGTA 3480
788-b1 > GGATCCCAATGCAAGGGCAACGGACTTCATGGATTGGCCAGAGTGGGGATCGAATCAAA 59
788pax1 > TGGATCCCAATGCAAGGGCAACGGACTTCATGGATTGGCCAGAGTGGGGATCGAATCAAA 315
788pax11 < TGGATCCCAATGCAAGGGCAACGGACTTCATGGATTGGCCAGAGTGGGGATCGAATCAAA 421
Consensus> TGGATCCCAATGCAAGGGCAACGGACTTCATGGATTGGCCAGAGTGGGGATCGAATCAAA 3540
788-b1 > CACTGATGCAGTTCTTCAATGATCGAGGTGCACTTCTTGCGGATAACTTTCGACAGGATA 119
788pax1 > CACTGATGCAGTTCTTCAATGATCGAGGTGCACTTCTtGCGGATAACTTTCGACAGGATA 375
788pax11 < C 422
Consensus> CACTGATGCAGTTCTTCAATGATCGAGGTGCACTTCTTGCGGATAACTTTCGACAGGATA 3600
788-b1 > CTTATGATTTTATTCTGGAGAATGTAGGGTCTTTCCATATCTAGGAAGCCAACATTTAAT 179
788pax1 > CTTAT 380
Consensus> CTTATGATTTTATTCTGGAGAATGTAGGGTCTTTCCATATCTAGGAAGCCAACATTTAAT 3660


788-b1 > AATGAGATATTGATCAATGTTTACAATGATGAATAATTGCCCACGAAATTCATGAATTTG 239
Consensus> AATGAGATATTGATCAATGTTTACAATGATGAATAATTGCCCACGAAATTCATGAATTTG 3720
788r < TAGT 4
788-b1 > GAGTTGAAGTAGTCAAACAAAGGTACTAATCTTCAGTATCTGAAGCTGACTTGAACTAGT 299
Consensus> GAGTTGAAGTAGTCAAACAAAGGTACTAATCTTCAGTATCTGAAGCTGACTTGAACTAGT 3780
788r < GATATCTAGCAGCTATCGCCAGGTCGCAGTAACGGTAGCTGTAGATCCGGAGCAGCCAGT 64
788-b1 > GATATCTAGCAGCTATCGCCAGGTCGCAGTAACGGTAGCTGTAGATCCGGAGCAGCCAGT 359
Consensus> GATATCTAGCAGCTATCGCCAGGTCGCAGTAACGGTAGCTGTAGATCCGGAGCAGCCAGT 3840
788r < GGGGACGGTGTAGGACCTGTTCGGATCGCTCTCCCAAGTCACCGATGAAGACCCACTCGA 124
788-b1 > GGGGACGGTGTAGGACCTGTTCGGATCGCTCTCCCAAGTCACCGATGAAGACCCACTCGA 419
Consensus> GGGGACGGTGTAGGACCTGTTCGGATCGCTCTCCCAAGTCACCGATGAAGACCCACTCGA 3900
788r < CTTTTCAAATAAACTTATACTCAAAAGACGAGCCAACGGGGATAGTCACCTCAACATACCA 184
788-b1 > CTTTTCAAATAAACTTATACTCAAAAGACGAGCCAACGGGGATAGTCACCTCAACATACCA 479
Consensus> CTTTTCAAATAAACTTATACTCAAAAGACGAGCCAACGGGGATAGTCACCTCAACATACCA 3960
788r < AAGTGGATTGGAAGAGGTATACTGATCTGCAGAAGAGCAATGGCATCATCGGTGTCCCA 244
788-b1 > AAGTGGATTGGAAGAGGTATACTGATCTGCAGAAGAGCAATGGCATCATCGGTGTCCCA 539
Consensus> AAGTGGATTGGAAGAGGTATACTGATCTGCAGAAGAGCAATGGCATCATCGGTGTCCCA 4020
788r < GCTACCCAACTGGCTGATTGAGCCTGAGAGGTAAACATTTTGGCCATATGAGGTGGTGAC 304
788-b1 > GCTACCCAACTGGCTGATTGAGCCTGAGAGGTAAACA 576
Consensus> GCTACCCAACTGGCTGATTGAGCCTGAGAGGTAAACATTTTGGCCATATGAGGTGGTGAC 4080
788r < GACCTCTTTGAACAGCACGGGTAAGGCAGTGGCCTGGGTACAGGTTGAAGTGGTACTGGC 364
Consensus> GACCTCTTTGAACAGCACGGGTAAGGCAGTGGCCTGGGTACAGGTTGAAGTGGTACTGGC 4140
788r < ACCGCCGGTACTTGAGGTCTTTGAACTGGTGCTCGTGGTTGACGCGCTTGTCGTAGTTCC 424
Consensus> ACCGCCGGTACTTGAGGTCTTTGAACTGGTGCTCGTGGTTGACGCGCTTGTCGTAGTTCC 4200
788r < GGAGTCTGAAGATGCACAGGGTCCACTGCTCTTCGCGGCAGAAGCAAGCATCAACACTCG 484
Consensus> GGAGTCTGAAGATGCACAGGGTCCACTGCTCTTCGCGGCAGAAGCAAGCATCAACACTCG 4260
788r < AGGCAAGCCAGAGACCATTGGGACAGCGATGTTGCCGCTCGAGTCGACTGTTATGGATGT 544
Consensus> AGGCAAGCCAGAGACCATTGGGACAGCGATGTTGCCGCTCGAGTCGACTGTTATGGATGT 4320
788r < ACATGTGTACATTTCCATGAGCTC 568
Consensus> ACATGTGTACATTTCCATGAGCTC 4344

*Appendix A2.3* *Sequence data of pPN1422.*

Spot sequencing of pPN1422 with pUC forward primer.

```
  1  TCTAGAGGAT CCCCGGGTAC CGAGCTCTTG GCATATGCCC CATCGGATCC
 51  WATGTTATTG ARAATCCCCT GTNTTGCAAC CGAAGTTTCC GATGCCAGSC
101  NAGTATCTAG GATACCTGTT GACCGGGGTG CTAGTTGAGG CGYGGCCAGT
151  ACCGCTGTAC CAATAGTCAG YGCACATAGY GCAGTCCAGA AACGAGGAGC
201  CATGATGAAA GATTGASCGT CGTTGCCTGG GCTKCAAAAG TAGCCTTGAT
251  AACCGGATMA GTCGTATGGT TTTYATGYTT CGACACGNTN TCTATCAGCT
301  CATATCGGNT TGATTACAAA TGTGAGTATA AAATTGCAGG CGNCGYGACA
351  AGGGGGTGGG GARRGTGTGA TGAAGCGGAG ATCGASGGGC TKTTGACTAG
401  NGACTYNNGA CGCAGGTGGT YTTGCTTATG GGAGNATCGC ATCTCTNANG
451  CCCCCAGCCA CAAANTTACT CTNCCCCCTC CACTMCAACT TCACTTCACT
501  TGGCGATC
```

Spot sequencing of pPN1422 with pUC reverse primer.

```
  1  TTAGATTCGG GTCACGCCGG AAGATATCCG CTGCTGCTGA GGCTCAAAGT
 51  GTAGGAGCTT CCCGAAGAGC CATAGTTCGA TAGGACTGTG ATGACTTGTN
101  AACCACTTGA CCCTTTCTTC ATCGCAATTG TGTGACTATC CTGGTAGAAT
151  CGTCCGTTCT AGGAACGGTT AGTCGATGCA TTGAATGGAT TTGTTGTAAG
201  ATCTGACCTT GGATGTTATG TAAGAAGAGT CTNAAGACAC TGCTAGAGCG
251  CGGATTTTGT TCGTCGTCGC AATGTGTTTG TACAGCTCAG CAGAGGTCGA
301  GTAACCGGAA AGSCATGTCG CTTCGCGATT AGCAGGATCA CTGCCTCCAC
351  TATAGTGCTG CTCTTGGCCA GCGTAGATGA TTGGAATACC ATCAGAGAGG
401  AACATGAAAG ATATGACATT CTTCCCCTGT GAGTAGTCGC TTGTGTWCCN
451  GATTTANGGT TAGCATGGTC GACTCGAGGA TTTAAATTAG CAAAAATACT
501  CACGAGGCAA AGCGAGGGTT ATCGGNGATT CTCAATGAAG TTTCCGAGCA
551  AAGTGGGTCT TTGGAGTCGG AAGCGACCGA GCTTGATCAT GTTGTNGAGG
601  TCGNTTATTC TACCACTGGA AGATCAAATG
```

## *Appendix A2.4* *Sequence data of pPN1421.*

Spot sequencing of pPN1421 with pUC forward primer.

```
  1  CTTGGGTCAM GTTTGTCATG GGCGGTGGTG GACTTCTCCT ACTTGTCTCA

 51  TTCGTTCTGG TTGAGGCGTA CGTTGCTAAA GAACCGATTT TCAATCTCCG
                                                      <----
101  TATTCTCCGA CGACCAAATG TTGCGACAAG CTACATCATC GGATTTCTTC
     ----------------
         pax14

151  AAATCACGGS TCAATTAGGA ATGATGTTCT CCATCCCACT CTACTTCCAG

201  GTCACCGAGG GCGCCTCGAC AACCACCTCG GGAGGTCATC TCATCCCAGC

251  TGTCGTCGGA AACACAGTCG GTGGCCTACT AGCCGGTGGA TTCATTCGCA

301  AAACAGGTCG ATACAAGATG CTCCTCGTTC TAGCCGGACT CATCGNTTCA

351  ATCTCATATT GTCCTCCTGG TACCTCTTCT GGAACGGNGA TATCGGATTC

401  TCGGAGTCAC TCTATATCTT CCCCGGANGG AGTCGGGCAC GGNCATTGNA

451  TCTTKCATCA GNCNTCATTG CCATGACCGC CMWCCTTGGC CCAGAGGGAA

501  TGGNCANTGG NAACAAGTTG GCTNCAANNC TTGTTAGCAA ATTTCGCNAT

551  GACCCCTGGC
```

Spot sequencing of pPN1422 with pUC reverse primer.

```
  1  CCAGGCGGGG ATTGTGATTG CTAGTCCCAG NACTAGTAGT CCTGACTGTG

 51  AGTGCTCCTG CTCGATTKTC ATKTGATGAA TGGSTGATGA GTAACGTAAT

101  AGATTACTAT ACCTGGACTC GTGACTCGGC ACTTGTTATG AAAACCTNGG

151  TCGATCTCTT CAAGAATGGK GACACGGATC TGCTCACTGK CATCGAAGAA

201  NACATCAGGT CCCAGGGCTA TATCCAGAGT GKNTCTAACC CATCCNGSGG

251  TCTCTCGNGK GGCGGTCTGG GCGAACCCAA GTKCAATGTT GACGAGACAT

301  CGTATWCTGG ATCCTGGGGN CGACCTCAAC GTGATGGNCC TGCATTGNGA

351  GCTACCGCCC TGGTCGCGTT CGGACAGTGG TTGATTGTAA GATGATTCCC

401  TYGATTAGNT GAGGAGTATC TATCTGACAG GCTCCAGAGT AACGGATACA

451  CATCGACGGG CACCGMCATC GNCTGGNCTA TTGTCCNCAA CGATCTTTCT

501  TATGTTGGCC AGCATTGGGC ACAGCTCT
```

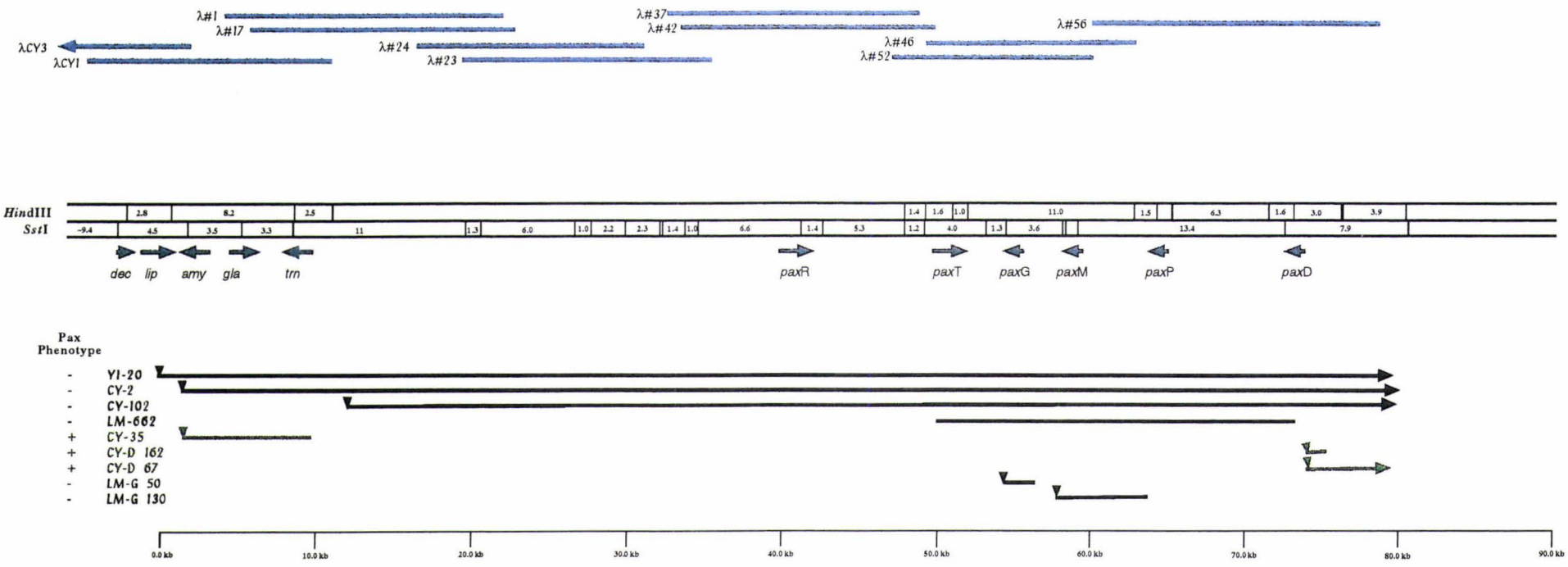
## *Appendix A2.5* *Sequence data of pPN1406.*

Spot sequencing of pPN1406 with pUC forward primer.

```
  1  GGCATTACGC TTGCTCACCA GGGGRAAACA TARAAAGGCA GAATCAATWC
 51  ATCGACATGT ATTGGAAANT CGAAAGAGAG NGATTGGACA CGACCATCCT
101  CATACACTCA CCACGGGGNA ATGACCTTGG CCTAGATCTA TTGAGACAAA
151  AGAAATATGA TGAAGCCGAC AAAATGTTTC GATGGGTTTT GGAAGCACGA
201  GAAAGGGTAC TTGGGCGCGA CCATCCCCAC ACGCTCATAA CTCTCAATGA
251  TGTTGGCCAC CTTCTCTCGA ATCAGGGAGA GCACGAGGAG GCAGAATCAA
301  TACAACGACA GGTATTAAAC ATGCGGGAGA GAATCCAAGG AGACGAGCAT
351  CCTGACACAC CCGTGACGGN CAATAACCTT GGCACAAATC TCTTGAGGCA
401  AGAGNAAATN TGAGGAGGCG GCNAGTGATG TTTCGGCGTG CGTTTTGGGT
451  TACGCCAAAM GGTGCTTGGG NGCACAAATC CCSCCACGCT TCCGTGTATG
501  AGGTGCC
```

Spot sequencing of pPN1406 with pUC reverse primer.

```
  1  GGGAAAGATT TTCCACCTTG ATCAAGAAGA CCTGTGATGG CGATAATCGC
 51  TGAGCCGAGA AGACCGTTCC GAGAAAATCG ASTTTTGCGA GTTTGGACAA
101  TGTGCTGTCT TCGTCGGAAG CTGTCTTCTG TGCTATGTTA AGCTTGATGA
151  CGACAAGGAG TGCAGCCAAA GCGACGAATG GAGCTTGGAT GAGGAACAGC
201  CTGGATTTNG TCAGCTGTTG ATTCCGGTCA AGGGTATTCG ATAGTCTTAC
251  CATCGCCATC CGATTGTATC GCTTAACCAC CCCCCGAGAG GACCACCGAG
301  ACTGCGTCCG AGAGTCATCG CGATATTGAC ATAGGCTCGY CATGTTGCCA
351  CTTCTCGCCT GGGTACGATA TCTGTGATTA TAACCGCAGC TGATCGTYCA
401  TCCGTGCCGN NNAYCACCCA TTCCGTTNAT GGCGCNNACC TAGAAATTGA
451  NCATCCCACA TACGAGTGTC AAAAACCACT AA
```




***Appendix A3.1*** *Chromosome walking*

# *Appendix*
# *A3.0 Chromosome Walking*

# *References*

Acklin W, Weibel F, Arigoni D (1977). Zur biosynthese von paspalin und verwandten metaboliten aus *Claviceps paspali*. Chimia 31: 63

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

Asch DK, Frederick G, Kinsey JA, Perkins DD (1992). Analysis of junction sequences resulting from integration at nonhomologous loci in *Neurospora crassa*. Genetics 130: 737-748

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

Barredo JL, Diez B, Alvarez E, Martin JF (1989). Large amplification of a 35-kb DNA fragment carrying two penicillin biosynthetic genes in high penicillin producing strains of *Penicillium chrysogenum*. Curr. Genet. 16: 453-359

Bennett JW, Papa KE (1988). The aflatoxigenic *Aspergillus* species. Adv. Plant Pathol. 6: 263-280

Beremand MN (1987). Isolation and characterization of mutants blocked in T-2 toxin biosynthesis. Appl. Environ. Microbiol. 53: 1855-1859

Bird D, Bradshaw R (1997). Gene targeting is locus dependent in the filamentous fungus *Aspergillus nidulans*. Mol. Gen. Genet. 255: 219-225

Bölker M, Böhnert HU, Braun KH, Görl J, Kahmann, R (1995). Tagging pathogenicity genes in *Ustilago maydis* by restriction enzyme-mediated integration (REMI). Mol. Gen. Genet. 248: 547-552

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

Bullock WO, Fernandez JM, Short JM (1987). XL1-Blue: a high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. BioTechniques 5: 376-378

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

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

Chappell J (1995). The biochemistry and molecular biology of isoprenoid metabolism. Plant Physiol. 107: 1-6

Cockrum PA, Culvenor CCJ, Edgar JA, Payne AL (1979). Chemically different tremorgenic mycotoxins in isolates of *Penicillium paxilli* from Australia and North America. J. Nat. Prod. 42: 534-536

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

Cole RJ, Kirksey JW, Wells JM (1974). A new tremorgenic metabolite from *Penicillium paxilli*. Can. J. Microbiol. 20: 1159-1162

Crowhurst RN, Hawthorne BT, Rikkerink EHA, Templeton MD (1991). Differentiation of *Fusarium solani* f. sp. *cucurbitae* races 1 and 2 by random amplification of polymorphic DNA. Curr. Genet. 20: 391-396

de Jesus AE, Gorst-Allman CP, Steyn PS, van Heerden FR, Vleggar R, Wessels PL, Hull WE (1983). Tremorogenic mycotoxins from *Penicillium crustosum*. Biosynthesis of Penitrem A. J. Chem. Soc.. Perkin Trans. 1863-1868

Diallinas G, Scazzocchio C (1989). A gene coding for the uric acid-xanthine permease of *Aspergillus nidulans*: inactivational cloning, characterization, and sequence of a *cis*-acting mutation. Genetics 122: 341-350

Dorner JW, Cole RJ, Cox RH, Cunfer BM (1984). Paspalitrem C, a new metabolite from sclerotia of *Claviceps paspali*. J. Agric. Food. Chem. 32: 1069-1071

Dower WJ, Miller JF, Ragsdale CW (1988). High efficiency transformation of *E. coli* by high voltage electroporation. Nucleic Acids Res. 16: 6127-6145

Feng GH, Chu FS, Leonard TJ (1992). Molecular cloning of genes related to aflatoxin biosynthesis by differential screening. Appl. Environ. Microbiol. 58: 455-460

Fierro F, Montenegro E, Gutiérrez S, Martin JF (1996). Mutants blocked in penicillin biosynthesis show a deletion of the entire penicillin gene cluster at a specific site within a conserved hexanucleotide sequence. Appl. Microbiol. Biotechnol. 44: 597-604

Frey M, Chomet P, Glawischnig E, Stettner C, Grün S, Winklmair A, Eisenreich W, Bacher A, Meeley RB, Briggs SP, Simcox K, Gierl A (1997). Analysis of a chemical plant defense mechanism in grasses. Science 277: 696-699

Gallagher RT, Hawkes AD, Steyn PS, Vleggaar R (1984). Tremorgenic neurotoxins from perennial ryegrass causing ryegrass staggers disorder of livestock: structure elucidation of lolitrem B. J. Chem. Soc. Chem. Commun 614-616

Gallagher RT, Latch GCM, Keogh RG (1980). The janthitrems: fluorescent tremorgenic toxins produced by *Penicillium janthinellum* isolates from ryegrass pastures. Appl. Environ. Microbiol. 39: 272-273

Gallagher RT, Wilson BJ (1978). Aflatrem, the tremorgenic mycotoxin from *Aspergillus flavus*. Mycopathologia 66: 183-185

Garthwaite I, Miles CO, Towers NR (1993). *Immunological detection of the indole diterpenoid tremorgenic mycotoxins.* In: Hume DE, Latch GCM, Easton HS (eds) Second International Symposium on *Acremonium*/Grass Interactions, Palmerston North, New Zealand. pp. 77-80

Gebler JC, Poulter CD (1992). Purification and characterisation of dimethylallyl trypothan synthase from *Claviceps purpurea*. Arch. Biochem. Biophys. 296: 308-313

Geiser DM, Arnold ML, Timberlake WE (1996). Wild chromosomal variants in *Aspergillus nidulans*. Curr. Genet. 29: 293-300

Hodges RL, Hodges DW, Goggans K, Xuei X, Skatrud P, McGilvray D (1994). Genetic modification of an echinocandin B-producing strain of *Aspergillus nidulans* to produce mutants blocked in sterigmatocystin biosynthesis. J. Ind. Microbiol. 13: 372-381

Hohn TM, Beremand PD (1989). Isolation and nucleotide sequence of a sesquiterpene cyclase gene from the trichothecene-producing fungus *Fusarium sporotrichioides*. Gene 79: 131-138

Hohn TM, McCormick SP, Desjardins AE (1993). Evidence for a gene cluster involving trichothecene-pathway biosynthetic genes in *Fusarium sporotrichioides*. Curr. Genet. 24: 291-295

Holmes DS, Quigley M (1981). A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. 114: 193-197

Hull EP, Green PM, Arst HN, Scazzocchio C (1989). Cloning and physical characterization of the L-proline catabolism gene cluster of *Aspergillus nidulans*. Mol. Microbiol. 3: 553-559

Ish-Horowitz D, Burke JF (1981). Rapid and efficient cosmid cloning. Nucleic Acids Res. 9: 2989-2998

Itoh Y (1994). Gene tagging in *Penicillium paxilli*. Tottori University.

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

Itoh Y, Scott B (1994). Heterologous and homologous plasmid integration at a spore-pigment locus in *Penicillium paxilli* generates large deletions. Curr. Genet. 26: 468-476

Itoh Y, Scott B (1997). Effect of dephosphorylation of linearised pAN7-1 and of addition of restriction enzyme on plasmid integration in *Penicillium paxilli*. Curr. Genet. 32: 147-152

Keller NP,Hohn TM (1997). Metabolic pathway gene clusters in filamentous fungi. Fungal Genet. Biol. 21: 17-29

Keller NP, Kantz NJ, Adams TH (1994). *Aspergillus nidulans verA* is required for production of the mycotoxin sterigmatocystin. Appl. Environ. Microbiol. 60: 1444-1450

Kimura N, Tsuge T (1993). Gene cluster involved in melanin biosynthesis of the filamentous fungus *Alternaria alternata*. J Bacteriol 175: 4427-4435

Kistler HC, Benny U (1992). Autonomously replicating plasmids and chromosome rearrangements during transformation of *Nectria haematococca*. Gene 117: 81-89

Laws I, Mantle PG (1989). Experimental constraints in the study of the biosynthesis of indole alkaloids in fungi. J. Gen. Microbiol. 135: 2679-2692

Liang P, Pardee AB (1992). Differential display of eukaryotic messenger RNA by means of the polymerse chain reaction. Science 257: 967-971

Lu S, Lyngholm L, Yang G, Bronson C, Yoder OC, Turgeon BG (1994). Tagged mutations at the *Tox1* locus of *Cochliobolus heterostrophus* by restriction enzyme-mediated integration. Proc. Natl. Acad. Sci. USA 91: 12649-12653

Mantle PG, Penn J (1989). A role for pacilline in the biosynthesis of indole-diterpenoid penitrem mycotoxins. J. Chem. Soc. Perkin Trans. 1: 1539-1540

Miles CO, Wilkins AL, Garthwaite I, Ede RM, Munday-Finch SC (1995). Immunochemical Techniques in Natrual products Chemistry: Isolation and structure determination of a novel indole-diterpenoid aided by TLC-ELISAgram. J. Org. Chem. 60: 6067-6069

Miller JH (1972). Experiments in Molecular Genetics. Cold Spring Harbor Laboratory Press, New York

Munday-Finch SC, Wilkins AL, Miles CO (1996). Isolation of paspaline B, an indole -diterpenoid from *Penicillium paxilli*. Phytochemisty 41: 327-332

Murray FR, Latch GCM, Scott DB (1992). Surrogate transformation of perennial ryegrass, *Lolium perenne*, using genetically modified *Acremonium* endophyte. Mol. Gen. Genet. 233: 1-9

Nozawa K, Udagawa S, Nakajima S, Kawai K (1987). Structures of two stereoisomers of a new type of indolediterpenoid related to the tremorgenic mycotoxin paxilline, from *Emericella desertorum* and *Emericella striata*. J. Chem. Soc., Chem. Commun. 1157-1159

Oliver RP, Roberts IN, Harling R, Kenyon L, Punt PJ, Dingemanse MA, van den Hondel CAMJJ (1987). Transformation of *Fulvia fulva*, a fungal pathogen of tomato, to hygromycin B resistance. Curr. Genet. 12: 231-233

Peberdy, JE (ed) (1987). Penicillium and Acremonium. Plenum Press, New York and London.

Perkins DD, Kinsey JA, Asch DK, Frederick GD (1993). Chromosome rearrangements recovered following transformation of *Neursopora crassa*. Genetics 134: 729-736

Proctor RH, Hohn TM, McCormick SP, Desjardins AE (1995). *Tri6* encodes an unusual zinc finger protein involved in regulation of trichothecene biosynthesis in *Fusarium sporotrichioides*. Appl Environ Microbiol 61: 1923-1930

Punt PJ, Oliver RP, Dingemanse MA, Pouwels PH, van den Hondel CAMJJ (1987). Transformation of *Aspergillus* based on the hygromycin B resistance marker from *Escherichia coli*. Gene 56: 117-124

Razanamparany V, Bégueret J (1988). Non-homologous integration of transforming vectors in the fungus *Podospora anserina*: sequences of junctions at the integration sites. Gene 74: 399-409

Sambrook J, Fritsch EF, Maniatis T (1989). Molecular cloning: a laboratory manual. New York, Cold Spring Harbor Laboratory Press.

Schiestl RH, Petes TD (1991). Integration of DNA fragments by illegitimate recombination in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA 88: 7585-7589

Seya H, Nozawa K, Udagawa S, Nakajima S, Kawai K (1986). Studies on fungal products. IX. Dethiosecoemestrin, a new metabolite related to emestrin, from *Emericella striata*. Chem. Pharm. Bull. 34: 2411-2416

Skory CD, Chang P-K, Cary J, Linz JE (1992). Isolation and characterization of a gene from *Aspergillus parasiticus* associated with the conversion of versicolorin A to sterigmatocystin in aflatoxin biosynthesis. Appl. Environ. Microbiol. 58: 3527-3537

Southern EM (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98: 503-517

Stryer L (1995). Biochemistry. New York, W.H. Freeman.

Thuring RW, Sanders PM. Borst P (1975). A freeze squeeze method for recovering DNA from agarose gel. Anal. Biochem 66: 213-220

Tilburn J, Roussel F, Scazzocchio C (1990). Insertional inactivation and cloning of the *wA* gene of *Aspergillus nidulans*. Genetics 126: 81-90

Tsai H-F, Wang H, Gebler JC, Poulter CD, Schardl CL (1995). The Claviceps purpurea gene encoding dimethylallyltryptophan synthase, the committed step for ergot alkaloid biosynthesis. Biochem. Biophy. Res. Commun. 216: 119-125

Vieira J and Messing J (1987). Production of single-stranded plasmid DNA. Methods Enzymol. 153: 3

Vollmer SJ, Yanofsky C (1986). Efficient cloning of genes of *Neurospora crassa*. Proc. Natl. Acad.Sci. USA 83: 4869-4873

Weedon CM. Mantle PG (1987). Paxilline biosynthesis by *Acremonium loliae*; a step towards defining the origin of lolitrem neurotoxins. Phytochemistry 26: 969-971

Weigel BJ, Burgett SG, Chen VJ, Skatrud PL, Frolik CA, Queener SW, Ingolia TD (1988). Cloning and expression in *Escherichia coli* of isopenicillin N synthetase genes from *Streptomyces lipmanii* and *Aspergillus nidulans*. J Bacteriol 170: 3817-3826

Wernars K, Goosen T, Swart K, van den Broek HWJ (1986). Genetic analysis of *Aspergillus nidulans* AmdS$^+$ transformants. Mol. Gen. Genet. 20: 312 -317

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research 18: 6531-6535

Yang G, Rose MS, Turgeon BG, Yoder OC (1996). A polyketide synthase is required for fungal virulence and production of the polyketide T-toxin. The Plant Cell 8: 2139-2150

Yelton MM, Hamer JE, Timberlake WE (1984). Transformation of *Aspergillus nidulans* by using a *trpC* plasmid. Proc. Natl. Acad. Sci. USA 81: 1470-1474

Yiang Y, Proteau P, Poulter D, Ferro-Novick S (1995). *BTS*1 encodes a geranylgeranyl diphosphate synthase in *Saccharomyces cerevisiae*. J. Biol. Chem. 270: 21793-21799

Yoder OC (1988). *Cochliobolus heterostrophus*, cause of southern corn leaf blight. Adv. Plant Pathol. 6: 93-112