

Research Paper

# Design and utility of oligonucleotide gene probes for fungal polyketide synthases

Thomas P. Nicholson <sup>a</sup>, Brian A.M. Rudd <sup>b</sup>, Mike Dawson <sup>b</sup>, Colin M. Lazarus <sup>c</sup>,  
Thomas J. Simpson <sup>a</sup>, Russell J. Cox <sup>a, \*</sup>

<sup>a</sup>School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK

<sup>b</sup>GlaxoSmithKline Medicines Research Centre, Gunnels Wood Road, Stevenage, Herts SG1 2NY, UK

<sup>c</sup>School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK

Received 10 October 2000; revisions requested: 17 November 2000 revisions received 29 November 2000; accepted 8 December 2000

First published online 22 December 2000

## Abstract

**Background:** Recent advances in the molecular biology of polyketide biosynthesis have allowed the engineering of polyketide synthases and the biological ('combinatorial') synthesis of novel polyketides. Additional structural diversity in these compounds could be expected if more diverse polyketide synthases (PKS) could be utilised. Fungal polyketides are highly variable in structure, reflecting a potentially wide range of differences in the structure and function of fungal PKS complexes. Relatively few fungal synthases have been investigated, perhaps because of a lack of suitable genetic techniques available for the isolation and manipulation of gene clusters from diverse hosts. We set out to devise a general method for the detection of specific PKS genes from fungi.

**Results:** We examined sequence data from known fungal and bacterial polyketide synthases as well as sequence data from bacterial, fungal and vertebrate fatty acid synthases in order to determine regions of high sequence conservation. Using individual domains such as  $\beta$ -ketoacylsynthases (KS),  $\beta$ -ketoreductases (KR) and methyltransferases (MeT) we determined specific short (ca 7 amino acid) sequences showing high conservation for particular functional domains (e.g. fungal KR domains involved in producing partially reduced metabolites; fungal KS domains involved in

the production of highly reduced metabolites etc.). Degenerate PCR primers were designed matching these regions of specific homology and the primers were used in PCR reactions with fungal genomic DNA from a number of known polyketide producing species. Products obtained from these reactions were sequenced and shown to be fragments from as-yet undiscovered PKS gene clusters. The fragments could be used in blotting experiments with either homologous or heterologous fungal genomic DNA.

**Conclusions:** A number of sequences are presented which have high utility for the discovery of novel fungal PKS gene clusters. The sequences appear to be specific for particular types of fungal polyketide (i.e. non-reduced, partially reduced or highly reduced KS domains). We have also developed primers suitable for amplifying segments of fungal genes encoding polyketide C-methyltransferase domains. Genomic fragments amplified using these specific primer sequences can be used in blotting experiments and have high potential as aids for the eventual cloning of new fungal PKS gene clusters. © 2001 Elsevier Science Ltd. All rights reserved.

**Keywords:** Polyketide synthase; Fungus; Squalestatin; Mycotoxin;  $\beta$ -Ketoacylsynthase; C-Methyltransferase

## 1. Introduction

Polyketide biosynthesis has been the focus of intense research over the past decade [1]. Advances in the understanding of the biosynthetic processes, gained through feeding experiments with labelled precursors led on to en-

zymatic experiments, and latterly and more profitably, investigations of the genes encoding polyketide synthases. The development of methods for the manipulation of these gene clusters has enabled the directed engineering of new synthases able to make new compounds. Because of the wide usage of polyketides as pharmaceuticals this methodology offers the ability to synthesise new drug candidates quickly, cleanly and in a combinatorial manner. So far the greatest success has come from manipulations of polyketide synthases obtained from the actinomycetes. In part this success stems from the development of useful tools

\* Correspondence: Russell J. Cox;  
E-mail: r.j.cox@bristol.ac.uk

for the genetic manipulation and discovery of genes in actinomycetes and other bacteria.

In seminal work focusing on the biosynthesis of actinorhodin **1** in *Streptomyces coelicolor* A3(2) [2], Hopwood and coworkers isolated a number of genes encoding proteins operating during the biosynthesis [3]. Genes known to function early during the biosynthesis of actinorhodin **1** (*actI* and *actIII*) were later found to encode the key components acyl carrier protein (ACP, *actI* ORF3),  $\beta$ -ketoacylsynthase ( $KS_{\alpha}$ , *actI* ORF1) and another protein highly homologous to  $KS_{\alpha}$  but with several active site mutations including glutamine rather than cysteine (designated here  $KS_{\beta}$ , *actI* ORF2) [4]. The act  $\beta$ -ketoreductase (KR, *actIII*) was also sequenced [5]. DNA fragments from the *act* cluster were used as probes in Southern blotting experiments to find other similar polyketide synthase (PKS) clusters in heterologous actinomycete hosts [6]. Such work led to the discovery of a number of Type II PKS gene clusters responsible for the biosynthesis of pharmacologically important compounds such as daunorubicin and daunomycin [7,8].

As a result of the success of the *act* probes in isolating novel PKS genes, domains from the *eryA* modular PKS gene (involved in erythromycin aglycone biosynthesis) were likewise used in an attempt to isolate modular PKS for a number of complex polyketide metabolites. The most notable success has come with the isolation of the PKS genes involved in the biosynthesis of the immunosuppressant macrolide rapamycin (*rap*) from *Streptomyces hygroscopicus*. This was achieved by cross-hybridisation of *eryA*-derived probes to the *rap* PKS genes in a *S. hygroscopicus* genomic DNA library [9]. More recently  $KS$ -derived probes from the rifamycin cluster have been used to screen a bacterial artificial chromosome library to detect genes encoding epothilone biosynthesis in the myxobacterium *Sorangium cellulosum* [10].

In general it appears that these heterologous probing approaches work well when probes are used to search for very similar types of PKS. For example the *act* probes work well for detecting other bacterial Type II PKS genes, but they do not hybridise well with modular PKS genes or even Type II bacterial fatty acid synthase (FAS) genes. Likewise the *ery* and other modular PKS probes work well for detecting modular PKS genes, but not for detecting Type II genes.

In fungi relatively few PKS gene clusters have been investigated [11]. The most well understood system is that responsible for the biosynthesis of 6-methylsalicylic acid (6MSA **2**). 6MSA synthase (MSAS) was isolated and purified from *Penicillium patulum* [12]. Antibodies raised against the purified protein were used to screen an *Escherichia coli* expression library of *P. patulum* DNA and a genomic clone was subsequently isolated. In this way the *P. patulum* MSAS gene was cloned and sequenced [13]. This gene was shown to encode a large multifunctional (Type I) protein of 1735 amino acids. The MSAS sequence

contained regions highly homologous to sequences encoding  $\beta$ -ketosynthase (KS), KR, dehydratase (DH), acyltransferase (AT) and ACP. Expression of the MSAS gene in the heterologous hosts *Aspergillus nidulans*, *S. coelicolor* [14], *E. coli* and *Saccharomyces cerevisiae* [15] led to production of 6MSA **2**. The availability of the MSAS gene sequence permitted DNA probing experiments in heterologous fungal systems. Such investigations led to the discovery of MSAS genes in the filamentous fungi *Aspergillus terreus* and *Aspergillus parasiticus* [16–18].

Other fungal PKS genes have also been investigated. Tetrahydroxynaphthalene **3** synthase (THNS) from *Colletotrichum lagenarium* was cloned from a cosmid library by complementation of an albino mutant and its activity proven by subsequent heterologous expression in *Aspergillus oryzae* [19]. Comparisons between null and producing mutants of the norsolorinic acid (NSA **4**) producers *A. nidulans* and *A. parasiticus* led to the eventual cloning of the NSA synthase gene (NSAS) from the aflatoxin biosynthetic cluster [16,20]. The *wA* gene responsible for the production of YWA1 **5**, a spore pigment, in *A. nidulans* has also been cloned, again by complementation of a null mutant [21]. The T-toxin synthase (TTS) of *Cochliobolus heterostrophus* was identified by comparison of two isogenic strains (race T and race O), that differ because race O does not produce T-toxin **6** [22]. Confirmation of PKS function was obtained by targeted gene disruption.

The lovastatin nonaketide synthase (LNKS) gene was isolated via mutagenesis and screening of *A. terreus* mutants for disruption of lovastatin production [23]. Mutants incapable of synthesising the nonaketide and diketide components were found to be missing polypeptides of  $\sim 250$  kDa and  $\sim 220$  kDa respectively. Purification of the LNKS protein and immunological screening of a LNKS containing cDNA expression library led to isolation of the LNKS gene and the lovastatin diketide synthase (*LDKS*) gene was then cloned by sequencing the surrounding region. Restoration of lovastatin biosynthesis by complementation with the cloned gene was used as confirmation of its function [24,25]. Very recently the synthase associated with the synthesis of fumonisin **8** in *Gibberella fujikuroi* was also cloned [26].

We have been seeking to devise a general and simple method for the cloning of diverse fungal PKS genes. Homologous hybridisation could be used for this purpose and in preliminary work in our laboratories we have investigated this approach [27]. Using sequence data available for the KS domains of (the then known) four fungal PKS genes we showed that the sequences could be grouped into MSAS-like and *wA*-like classes. Degenerate PCR primers specific for each class were designed, and used in PCR reactions with genomic DNA. The PCR products were then used in homologous and heterologous Southern blotting experiments. The results showed that the division into two types of PKS held true and that the probes could be used to identify new (albeit very

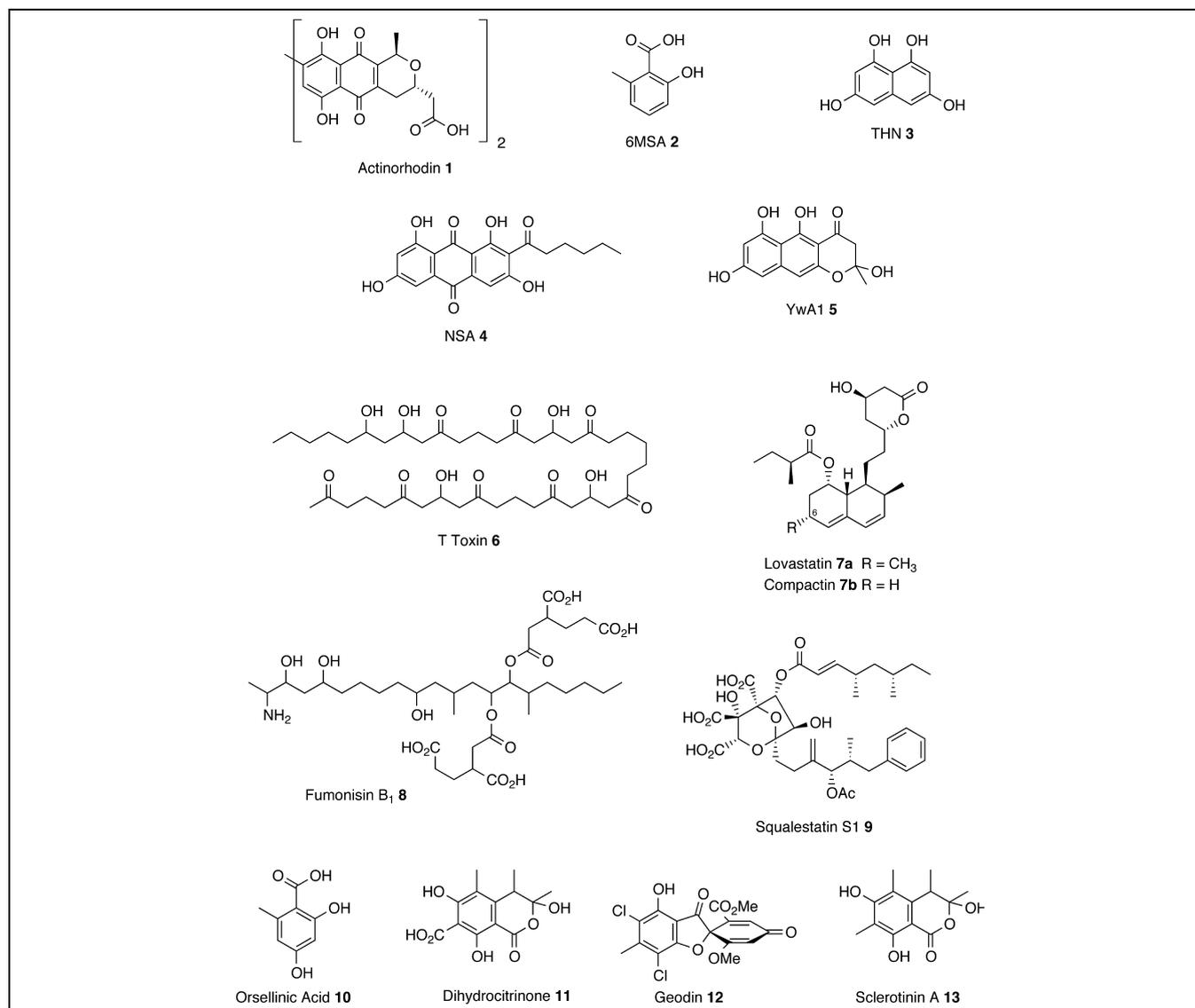
closely related or homologous) PKS genes. Importantly we also showed that these probes could distinguish between PKS genes of differing subtype (i.e. *MSAS*-like and *wA*-like) within a single host.

The results from our previous work indicate the possibility of designing subclass specific PCR primers for use in fungal PKS cloning strategies involving homologous probe generation. Here we describe the extension of this idea to three fungal PKS KS subtypes and the use of new sequence data for the design of probes based on KR and C-methyltransferase (CMeT) domains.

## 2. Results

Our previous results have indicated that different classes of fungal PKS genes can be distinguished by DNA probes. Although we used DNA probes generated from PCR

primers based on only four sequences, it appeared that the discrimination into subclasses could be based on the extent of reduction of the polyketide being produced [27]. This correlates with the structures and known biosynthetic origins of fungal polyketides. For example there is a large class of fungal polyketides, represented by 1,3,6,8-tetrahydroxy naphthalene **3**, NSA **4**, YwA1 **5** and orsellinic acid **10** where there have been no reductive steps during biosynthesis. Another class, exemplified by 6MSA **2** shows partial reduction (i.e. a single round of KR and DH for 6MSA **2**). A third class, containing compounds such as T-toxin **6**, lovastatin **7a** and squalestatin **9** show much higher levels of reductive modification (for structures see Scheme 1). For the purpose of this discussion it is convenient to refer to these structural types as non-reduced (NR), partially reduced (PR) and highly reduced (HR). Because fungal hosts often contain multiple PKS genes it would be useful to have probes that can distinguish between



Scheme 1.

Table 1  
Gene sequences considered during the design of PKS selective PCR primers

| Organism                          | gene             | Synthase component | Metabolite       | Type        | Database accession number |
|-----------------------------------|------------------|--------------------|------------------|-------------|---------------------------|
| <i>A. parasiticus</i>             | <i>pksL1</i>     | NAS                | NA 4             | NR PKS      | sw:PKS1_ASPPA             |
| <i>A. nidulans</i>                | <i>pksST</i>     | NAS                | NA 4             | NR PKS      | sw:STCA_EMENI             |
| <i>A. nidulans</i>                | <i>wA</i>        | wAS                | YWA1 5           | NR PKS      | em:X65866                 |
| <i>C. lagenarium</i>              | <i>pksI</i>      | THNS               | THN 3            | NR PKS      | em:D836                   |
| <i>P. patulum</i>                 | <i>MSAS</i>      | MSAS               | 6MSA 2           | PR PKS      | sw:MSAS_PENPA             |
| <i>A. terreus</i>                 | <i>MSAS</i>      | MSAS               | 6MSA 2           | PR PKS      | em:ATU31329               |
| <i>A. parasiticus</i>             | <i>pksL2</i>     | MSAS               | 6MSA 2           | PR PKS      | em:APU52151               |
| <i>Penicillium griseofulvum</i>   | <i>PKS2</i>      | MSAS               | 6MSA 2           | PR PKS      | em:PGU89769               |
| <i>A. terreus</i>                 | <i>lovF</i>      | LDKS               | lov diketide 7   | HR PKS      | em:AF141925               |
| <i>A. terreus</i>                 | <i>lovB</i>      | LNKS               | lov nonaketide 7 | HR PKS      | em:AF151722               |
| <i>C. heterostrophus</i>          | <i>pksI</i>      | TTS                | T-toxin 6        | HR PKS      | em:CHU68040               |
| <i>G. fujikuroi</i>               | <i>fum5</i>      | FUMS               | fumonisin 8      | HR PKS      | em:AF155773               |
| <i>Saccharopolyspora erythrae</i> | <i>eryA ORF1</i> | DEBS               | 6dEB             | modular PKS | sw:ERY1_SACER             |
| <i>S. erythrae</i>                | <i>eryA ORF3</i> | DEBS               | 6dEB             | modular PKS | sw:ERY3-SACER             |
| <i>S. hygrosopicus</i>            | <i>rapA mod1</i> | RAPS               | rapamycin        | modular PKS | em:SHGCPIR                |
| <i>S. hygrosopicus</i>            | <i>rapB mod5</i> | RAPS               | rapamycin        | modular PKS | em:SHGCPIR                |
| <i>S. coelicolor</i>              | <i>act1 ORF1</i> | Act KS $\alpha$    | actinorhodin 1   | Type II PKS | sw:KASA_STRCO             |
| <i>E. coli</i>                    | <i>fabB</i>      | KASI               | fatty acids      | Type II FAS | sw:FABB_ECOLI             |
| <i>Rattus norvegicus</i>          | <i>Fas</i>       | FAS                | fatty acids      | Type I FAS  | sw:FAS_RAT                |
| <i>P. patulum</i>                 | <i>FAS2</i>      | FAS                | fatty acids      | Type I FAS  | sw:FAS2_PENPA             |
| <i>A. nidulans</i>                | <i>FASA</i>      | FAS                | NA FAS           | Type I FAS  | em:ENU75347               |
| <i>E. coli</i>                    | <i>fabG</i>      | KR                 | fatty acids      | Type II FAS | sw:FABG_ECOLI             |
| <i>S. coelicolor</i>              | <i>Act III</i>   | KR                 | actinorhodin     | Type II PKS | sw:ACT3_STRCO             |
| <i>Streptomyces griseus</i>       | <i>Gris ORF5</i> | KR                 | griseusin        | Type II PKS | em:SGKSACPG               |
| <i>Streptomyces violaceoruber</i> | <i>Gra ORF5</i>  | LR                 | granaticin       | Type II PKS | em:SVPKS                  |
| <i>A. parasiticus</i>             | <i>ver1</i>      | PhR                | aflatoxin        | post PKS    | em:APVER1A                |
| <i>A. parasiticus</i>             | <i>norA</i>      | PhR                | aflatoxin        | post PKS    | em:AP24698                |
| <i>C. lagenarium</i>              | <i>thr1</i>      | PhR                | melanin          | post PKS    | em:CLTHR1                 |
| <i>M. grisea</i>                  | <i>phnr</i>      | PhR                | melenin          | post PKS    | em:MGPORE                 |
| <i>S. peuceitius</i>              | <i>dnrK</i>      | O-MT               | daunorubicin     | post PKS    | sw:CM4T_STRPE             |
| <i>S. glaucesens</i>              | <i>tcnN</i>      | O-MT               | tetracenomycin   | post PKS    | sw:TCMN_STRGA             |
| <i>A. parasiticus</i>             | <i>omtA</i>      | O-MT               | aflatoxin        | post PKS    | sw:OMTA_ASPPA             |
| <i>R. norvegicus</i>              | <i>HMT</i>       | N-MT               | N-methyl his     | regulatory  | sw:HMT_RAT                |
| <i>R. norvegicus</i>              | <i>GLMT</i>      | N-MT               | N-methyl gly     | regulatory  | sw:GLMT_RAT               |
| <i>S. lipmanii</i>                | <i>Pur5</i>      | N-MT               | puromycin        | antibiotic  | em:Q53742                 |

PhR, Phenol reductase; em, EMBL database; sw, SWISSPROT database; other abbreviations explained in the text.

PKS gene classes. We therefore set out to analyse the available fungal PKS gene sequences with the aim of producing probes specific for NR, PR and HR PKS genes.

### 2.1. Design of degenerate oligonucleotide PCR primers

Our overall strategy was to design PCR primers, specific for certain classes of PKS, which could be used to generate a homologous probe (or probes) from any given fungal genomic DNA. The DNA and protein sequences we chose are summarised in Table 1. We used all the available sequence data on fungal PKS clusters, except FUMS which was not available at the outset of this study [26]. In order to avoid the possibility of mis-priming with endogenous FAS genes we also considered representative FAS sequences from fungi, bacteria and vertebrates. In order to compare our sequences with known bacterial PKS clusters we also included representative sequences from modular (*ery* and *rap*) and Type II (*act*) PKS clusters.

Our previous work concentrated on the use of probes designed with potential homology to KS domains alone. In this study we also wished to examine the possibility of using sequence data from KR domains. We used sequences for fungal PKS KR, bacterial modular KR, bacterial Type II PKS KR as well as fungal, vertebrate and bacterial FAS KR domains. In addition we used sequences from known post-PKS reductase enzymes, some of which such as the *C. lagenarium* THN reductase (*thr1*) are more properly regarded as phenol reductases.

One major difference between fungal PKS and bacterial PKS systems is the introduction of methionine-derived methyl branches by specialised CMeT domains. The use of CMeT-based probes could give additional specificity to any cloning approach. We therefore used the only available PKS CMeT domain sequences, from the *A. terreus* LDKS and LNKS PKS genes as the basis for our primers. In order to discriminate against general O- and N-methyltransferases, sequences for post-PKS enzymes involved in

Table 2  
Strains and clones used in this study

| <i>E. coli</i> Strain  | Genotype  | Reference/Source      |
|--|---|-----------------------|
| DH5 $\alpha$ <sup>®</sup>  | $\phi$ 80dlacZ $\Delta$ M15, <i>recA1</i> , <i>endA1</i> , <i>gyrA96</i> , <i>thi-1</i> , <i>hsdR17</i> ( <i>r<sub>K</sub>-</i> , <i>m<sub>K</sub>+</i> ), <i>supE44</i> , <i>relA1</i> , <i>deoR</i> , $\Delta$ ( <i>lacZYA-argF</i> )U169                                 | [38]                  |
| JM109  | <i>endA1</i> , <i>recA1</i> , <i>gyrA96</i> , <i>thi</i> , <i>hsdR17</i> ( <i>r<sub>K</sub>-</i> , <i>m<sub>K</sub>+</i> ), <i>relA1</i> , <i>supE44</i> $\Delta$ ( <i>lac-proAB</i> ), [ <i>F'</i> <i>traD36</i> , <i>proAB</i> , <i>laqI</i> <sup>q</sup> Z $\Delta$ M15] | Promega               |
| Fungal species   | Metabolite produced   | Reference/Source      |
| <i>Phoma</i> sp. C2932 (Strain PA13)   | squalestatin <b>S<sub>1</sub> 9</b>   | GlaxoSmithKline       |
| Unidentified species Merck strain MF5453                                       | squalestatin <b>S<sub>1</sub> 9</b>   | ATCC                  |
| <i>P. patulum</i> IMI92273 ATCC 48325  | 6MSA (patulin) <b>2</b>   | ATCC                  |
| <i>A. terreus</i> ATCC 20542   | lovastatin <b>7a</b>  | ATCC                  |
| <i>P. citrinum</i> NRRL 8028   | compactin <b>7b</b>   | NRRL                  |
| <i>C. heterostrophus</i> strain C4 ( <i>ToxI</i> <sup>+</sup> ; <i>MAT-2</i> ) | T-toxin <b>6</b>  | Prof. Olen Yoder      |
| Clone  | Utility   | Reference/Source      |
| pBluescript SK   | pUC19-derived cloning vector  | Stratagene            |
| pGEM-T   | T-overhang cloning vector   | Promega               |
| pMSAS-2  | Vector containing <i>P. patulum</i> <i>MSAS</i> gene  | Dr. Colin Lazarus     |
| pTPKS-100  | Vector containing <i>A. terreus</i> <i>LNKS</i> gene  | Prof. C.R. Hutchinson |
| p627PKSCH-28   | Vector containing <i>C. heterostrophus</i> <i>pkS1</i> gene   | Prof. Olen Yoder      |
| pERY9KS1/2, KR1/2)   | Vectors containing <i>S. erythrae</i> <i>eryA ORF1</i> fragments  | Prof. Peter Leadlay   |
| pRJC II 04F/pALM I 32E   | Vectors containing <i>S. coelicolor</i> <i>actI/lactIII</i> genes   | Dr. Anne-lise Matharu |

the daunorubicin, tetracenomyacin and aflatoxin pathways as well as some more general *N*-methyltransferase sequences were considered.

## 2.2. Target organisms

The organisms used in this study are summarised in Table 2. *P. patulum* is known to produce 6MSA **2**, and sequence data is available for both *MSAS* and the Type I fungal FAS from this organism. We also sought to test our methodology using other fungal species with known PKS genes, such as the lovastatin producer *A. terreus* where the *LNKS* and *LDKS* genes have been sequenced; the compactin **7b** producer *Penicillium citrinum*, which should contain a very similar gene cluster (encoding a putative compactin synthase) which has not been investigated previously; and *C. heterostrophus*, the T-toxin producer. We also wished to examine fungi which produce interesting polyketides of known pharmacological activity but where no prior genetic information was available. For this purpose we chose the squalestatin **S<sub>1</sub> 9** producing strains *Phoma* sp. C2932 and the unidentified Merck strain MF5453.

Squalestatin **S<sub>1</sub> 9** is interesting because of its structural features: it is composed of two HR polyketide chains; the main chain is a hexaketide and the appended side-chain is a tetraketide. This indicates the presence of (at least) two PKS genes, by analogy to the biosynthesis of lovastatin **7a** by *A. terreus*. The starter unit of the main chain is derived from benzoate and both chains possess methyl branches

derived from methionine. It is therefore a fair assumption that the SQ synthase genes (putative squalene hexaketide synthase, SQHKS, and squalestatin tetraketide synthase, SQTKS) should possess CMeT domains similar to those found in LNKS and LDKS.

## 2.3. KS primers

### 2.3.1. NR/PR PKS selective primers

Our previous work described the design of degenerate KS primers using NR and PR sequences (*THNS*, *wA* and *MSAS* sequences respectively) [27]. These primer pairs were designated LC1/LC2c (NR) and LC3/LC5c (PR). From sequence alignment, the size of PCR products likely to be obtained from genomic DNA with the NR and PR primer pairs can be estimated at 720 and 680 bp respectively (assuming no introns). Genomic DNA was prepared from our selected fungal species and used in standard PCR reactions using the LC1/LC2c and LC3/LC5c primer pairs. PCR reactions were performed at constant primer concentrations, but in order to optimise product formation the reactions were run at varied genomic DNA template and Mg<sup>2+</sup> concentrations (see Fig. 1 for a representative example with MF5453).

PCR products of approximately the predicted length were obtained for the NR and PR primer pairs with both the MF5453 strain and *Phoma* sp. templates. These were cloned into *E. coli*. To allow for the possibility that individual PCR bands could contain multiple heterologous sequences, plasmid minipreps were obtained for at least 12

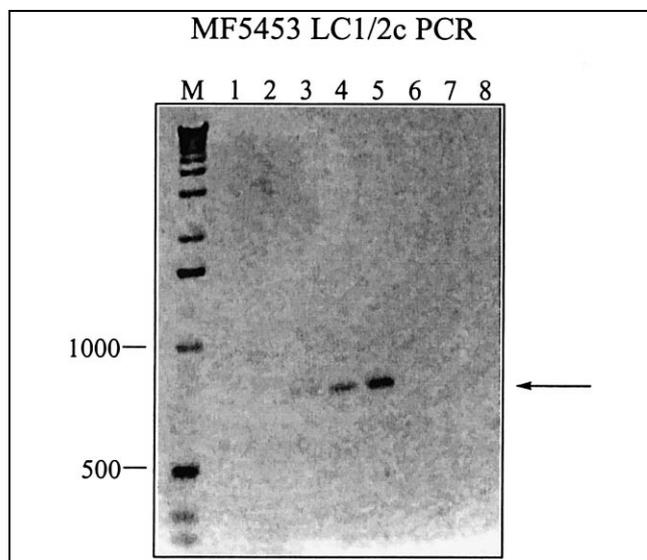


Fig. 1. Detection of NR PKS KS fragments. LC1/2c primer PCR with MF5453 strain. PCR products of approximately 750 bp, indicated by arrow, were obtained with genomic DNA template from MF5453 strain. Thermal cycling parameters were: (98°C, 5')×1, (94°C, 1'; 55°C, 1'; 72°C, 1')×34, (72°C, 10'). PCR primer concentrations were used at the following concentrations: LC1 3.2 µM, LC2c 0.4 µM. Template volume used (~10 ng µl<sup>-1</sup>): lanes 1–5, 5.0 µl; 6, 2.5 µl; 7, 10.0 µl; 8, 0.0 µl. MgCl<sub>2</sub> volume used (25 mM): lanes 1, 6–8, 0.0 µl; 2, 0.5 µl; 3, 1.0 µl; 4, 2.0 µl; 5, 7.0 µl. Reaction products were fractionated on a 1.5% agarose gel in TAE buffer at 100 mA, 200 V. M, 1 kb DNA ladder size markers, indicated in bp (Gibco BRL).

clones of each PCR product and each clone was sequenced. The sequence data revealed that each primer pair had amplified a single sequence from each species: Mfpk1 (LC1/2c) and Mfpk2 (LC3/5c) from strain MF5453 and Phpks1 (LC1/2c) and Phpks2 (LC3/5c) from *Phoma* sp.

Comparison of products derived from the two different genomic DNA templates with the same primer pair revealed that they were highly homologous (Fig. 2). The MF5453 and *Phoma* sp. NR products showed 98.9% nucleotide sequence and 97.5% deduced amino acid identity, whilst the PR products showed 85.7% nucleotide identity with 85.5% and 82.4% amino acid similarity and identity. This suggests that the products derived from the two different templates are essentially the same, or at least encode proteins that are functional homologues. BLAST sequence similarity searches of the EMBL database with the MF5453 and *Phoma* sp. product sequences indicated that they are highly homologous to various fungal and bacterial Type I PKS genes. The two sequences amplified using the NR primer pair show the highest homology to other NR class polyketide synthases such as THNS (88% similarity, 83% identity of translated polypeptide to *C. lagenarium* THNS) while the PR primed products are most homologous to MSAS (60–75% similarity/identity to *P. patulum* MSAS). Cloning and sequencing of additional bands observed in PR PCR product reactions did not give rise to any other PKS-like DNA sequences.

### 2.3.2. HR PKS selective primers

Using the protein sequences of LNKS from *A. terreus* and TTS from *C. heterostrophus*, we designed a third set of KS primers (Fig. 3). Regions of amino acid sequence conserved within the HR PKS subclass (but not between it and other types of PKS or related genes) were first identified. Deduced amino acid sequences from the KS domains of known fungal PKS genes were aligned, along with fungal FAS, bacterial FAS/PKS and mammalian FAS sequences, using the PileUp sequence alignment program. The sequences of MF5453 and *Phoma* sp. genomic

| A            |  |
|--------------|--|
| <i>Phoma</i> | PRFFNMSPRE AYQTDPMQRM ALTTAYEALE MSGYVNPRTA STRLDRIGTF YGQTSDDWRE INAAQEVDTY 70  |
| MF5453       | .....  |
| <i>Phoma</i> | FITGGVRAFG PGKINYHFGF SGPSLNDITA CSSSAAALQI ACTSLWAKDC DTAVVGGGLSC MNPDPFSGL 140 |
| MF5453       | .....V.....F.....T.....  |
| <i>Phoma</i> | SRGQFLSKTG PCATFDNGTD GYCRADGCAS VIVKRLDDAI ADKDNVLAVI LGTATNHSAD AISITHPHGP 210 |
| MF5453       | .....A.....  |
| <i>Phoma</i> | TQSILSSAIL DEAGVDPLDV DYVEMHGTGT 240   |
| MF5453       | .....D.....  |
| B            |  |
| <i>Phoma</i> | AEQMAPHQKL GLEVTWEALQ NPGINPKSLA GSdTAVYIGV HSDDYSRLLL EDIPNIEAWM GIGTTAHGIP 70  |
| MF5453       | .....D.....R.....E DA.....D.....R.....E.....S                                    |
| <i>Phoma</i> | NRISYHLNLM GFSAAVDAAC ASSLVAVNSG QQAILAGKSR VAIVGGVNVV LSPALFHMLG AAGALSPDGV 140 |
| MF5453       | .....D.....I.....H.....G.....VS.....G.....R.....G.....IK.....                    |
| <i>Phoma</i> | CLSPDDNARG YARGEGAAIL VLKKLSHAML DGDQVLVTIK GTAVAQDGKT NGIMAPNGKA QELVAWKALQ 210 |
| MF5453       | .....D.....C.....I.....R.....II.....H.....A.....R.....R.....K                    |
| <i>Phoma</i> | QANIDPLTVG YIEAHATFL 229   |
| MF5453       | .....R.....S   |

Fig. 2. Predicted amino acid sequence for PCR products obtained using (A) LC1/2c (NR) and (B) LC 3/5c (PR) primer pairs with MF5453 and *Phoma* sp. genomic templates. Sequences aligned using GAP sequence alignment program (Wisconsin Package, 10.0, GCG), with program default parameters. Sequences shown in the LC1→LC2c, LC3→LC5c direction. Full *Phoma* sp. sequence given, MF5453 sequence identical except where indicated by an alternative amino acid residue.

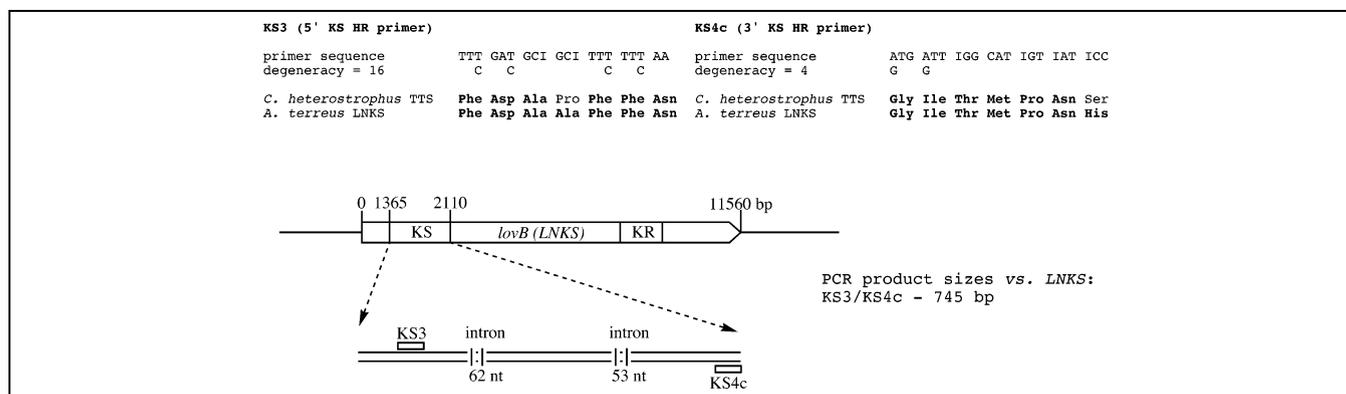


Fig. 3. Design of degenerate HR KS PCR primers. Primer sequence and design: Degenerate primer sequence and amino acid sequence of fungal PKS subclass upon which they are based. I = inosine, c = reverse complement nucleotide sequence. Bold text indicates bias to particular sequence conservation motif. Accession numbers for sequences are shown in Table 1.

template derived NR and PR PCR products were also included to ensure that the overlap between them and the HR PKS primer sequences was minimised. Close comparison of the aligned fungal PKS sequences revealed regions, within the HR PKS KS domains, of conserved amino acid sequence in different positions relative to those upon which the NR and PR primers were based.

The same general active site flanking regions were well conserved as observed for NR and PR sequences. Degenerate primers (KS3/KS4c) based on the conserved HR PKS sequence specific sections of this region were designed (Fig. 3). Care was taken to avoid known intron positions, and a PCR product size of ca 745 bp was predicted from primer binding positions relative to *LNKS*.

Control PCR reactions were run for the HR primer pair against pTPKS-100 (*LNKS*) and p627PKSCH-28 (*TTS*) plasmid templates. Gel electrophoresis of the resulting PCR product solutions revealed single bands of 700–750 bp for both templates (data not shown). These bands were cloned and sequenced and, as expected, the sequences were identical to those of the *LNKS* and *TTS* genes.

Gene fragment amplification reactions were run for genomic DNA templates from *A. terreus*, *C. heterostrophus*, unidentified Merck strain MF5453, *Phoma* sp. C2932 and *P. citrinum* with the HR KS primer pair (KS3/4c) (Fig. 4A for *A. terreus* results). The *LNKS* and *TTS* containing genomic DNA was used to ensure that the primers amplified DNA from the genes upon which they were based in a

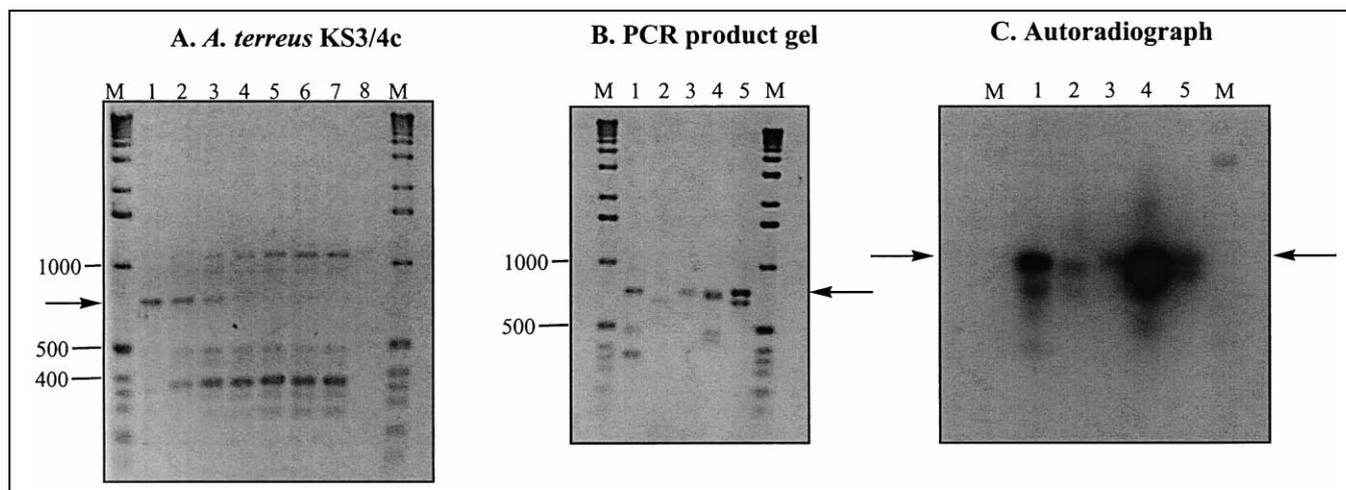


Fig. 4. Detection of HR PKS KS fragments. (A) PCR products of approximately 750 bp, indicated by arrows, were obtained with genomic template DNA from *A. terreus*. Thermal cycling parameters were: (98°C, 5') $\times$ 1, (94°C, 1'; 50°C, 1'; 72°C, 1') $\times$ 34, (72°C, 10'). PCR primers were used at the following concentrations: KS3, lane 1, 0.5  $\mu$ M; 2, 1.0  $\mu$ M; 3, 2.0  $\mu$ M; 4, 3.0  $\mu$ M; 5, 4.0  $\mu$ M; 6, 6.0  $\mu$ M; 7, 8.0  $\mu$ M; KS4c 1.0  $\mu$ M. Magnesium ion (MgCl<sub>2</sub>) concentration of 5 mM and approximately 5 ng of template DNA were used in all reactions except those in lane 8 (0 ng). Reaction products were fractionated on a 1.5% agarose gel in TAE buffer at 100 mA, 200 V. M, 1 kb DNA ladder size markers, shown in bp (Gibco BRL). (B) Products obtained by KS3/4c-primed PCR, with genomic template DNA from several filamentous fungi, were blotted from the agarose gel shown onto a nylon membrane and probed with <sup>32</sup>P-labelled Mf<sub>pk</sub>s3 KS PCR product. Hybridised bands in the 700–750 bp size range are indicated by arrows. Multiple PCR products were obtained with thermal cycling parameters as follows: (98°C, 5') $\times$ 1, (94°C, 1'; 50°C, 1'; 72°C, 1') $\times$ 34, (72°C, 10'). Primer concentrations of 1.0  $\mu$ M, 5 ng genomic template and 5 mM Mg<sup>2+</sup> were used in all the reactions. Genomic template DNA from the following organisms was used: lane 1, *A. terreus*; 2, *C. heterostrophus*; 3, *Phoma* sp.; 4, MF5453; 5, *P. citrinum*. Reaction products were fractionated on a 1.5% agarose gel in TAE buffer at 100 mA, 200 V. M, 1 kb DNA ladder size markers (Gibco BRL).



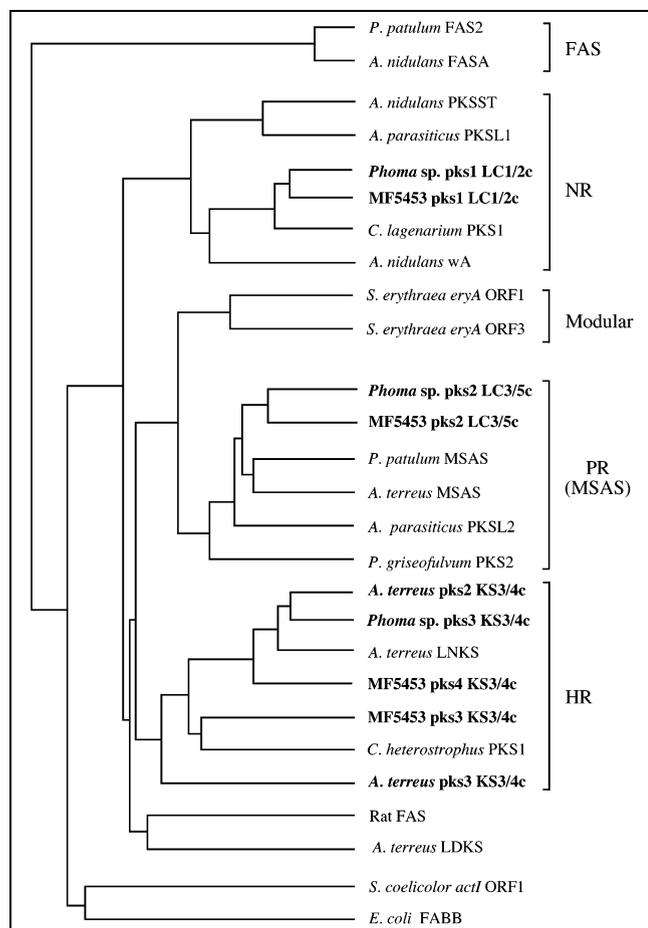


Fig. 7. TreePlot showing comparison of novel KS3/4c PCR product deduced amino acid sequences with those of known fungal PKS KS domains. PileUp run using the default specifications from Wisconsin package version 10.0, via GCG. Length of the lines is inversely proportional to the degree of sequence homology. Novel *A. terreus*, MF5453 and *Phoma* sp. KS3/4c sequences highlighted in bold type.

erable homology (>70%) to LNKS and *A. terreus* HR pks2 KS.

Multiple sequence analysis shows that the novel PKS KS products can be grouped with the known HR PKSs (Fig. 7). This indicates that neither of the novel *A. terreus* HR PKS KS products is derived from the *LDKS* gene and that the single condensation catalysing *LDKS* protein shows considerable differences in sequence to the iterative LNKS and TTS.

#### 2.4. CMeT primers

Methyl transfer from *S*-adenosylmethionine (SAM) to either nitrogen, oxygen or carbon atoms is catalysed by methyltransferases (MeT). The various aspects of the role of methylation in primary metabolism have been discussed elsewhere [29]. Methylation, usually of hydroxyl or carboxyl groups, is also employed in secondary metabolic biosynthetic pathways, with SAM as the methyl group donor. In fungal PKS systems *O*- and *N*-methylations

are catalysed by enzymes separate from the PKS, occurring towards the end of the biosynthetic pathway. For example the MeT responsible for the conversion of sterigmatocystin to *O*-methyl sterigmatocystin during aflatoxin B<sub>1</sub> biosynthesis has been purified from *A. parasiticus* [30]. Immunological screening was then used to isolate the corresponding full length cDNA transcript of the gene *omtA*, which has subsequently been used as a probe for cloning *OMeT* genes from other aflatoxin producers [31].

In contrast to *O*- and *N*-methylation, *C*-methylation in fungal PKS biosynthetic pathways takes place during polyketide chain formation, with the *CMeT* present as a well defined catalytic domain in the PKS itself. For example the LNKS and LDKS Type I PKS responsible for lovastatin **7a** biosynthesis in *A. terreus* both contain *CMeT* domains catalysing *C*-methylation at specific points on their respective nascent polyketide chains. Other fungal polyketides, including the squalestatins, employ *C*-methylation during polyketide assembly. The PKS involved in the biosynthesis of such metabolites are also therefore presumed to possess internal *CMeT* domains.

Regions of sequence similarity within a wide range of SAM dependent MeTs have been reported by Kagan and Clarke [32]. These workers identified three conserved sequence motifs, designated Motifs I, II and III, found in the same order and separated by comparable intervals in a variety of non-DNA MeTs. Despite low overall sequence homology between MeT domains from different sources Kagan and Clarke proposed that these regions of sequence similarity would be reflected in the 3D structures of such methylation reaction catalysing enzymes. Motifs I and III, denoted by the consensus amino acid sequences (L/I/V)-(V/L)(E/D)(V/I)G(C/G)G(P/T)G and LL(K/R)PGG(L/I/R)(I/L)(V/I/F/L)(L/I) respectively, were found in all known MeT domains, with Motif II, (G/P)(T/Q)(A/Y/F)DA-(Y/V/I)(I/F)(L/V/C), conserved in approximately 80% of the enzymes analysed. Both Motifs I and III have been identified in the *CMeT* domains of the LNKS and LDKS PKSs from *A. terreus*, although Motif II is absent. Motif I has been shown to be involved in the binding of SAM in MeT enzymes, represented by the  $\beta$ 1 sheet/ $\alpha$ 4 helix pocket region of the catechol *OMeT* 3D protein structure [33]. Motif III has been attributed to either the active site or magnesium cofactor binding site regions of MeT enzymes and is represented by the  $\beta$ 5 sheet region of the catechol *OMeT* 3D protein structure.

LDKS and LNKS gene and deduced amino acid sequences from *A. terreus* provided a means of analysing fungal PKS *CMeT* domains for regions of conserved sequence with a view to designing domain specific PCR primers. The deduced amino acid sequences of a number of small molecule MeT enzymes and domains were compared. Close comparison of the LDKS and LNKS *CMeT* domains revealed two regions of high amino acid sequence conservation approximately 100 amino acid residues apart,

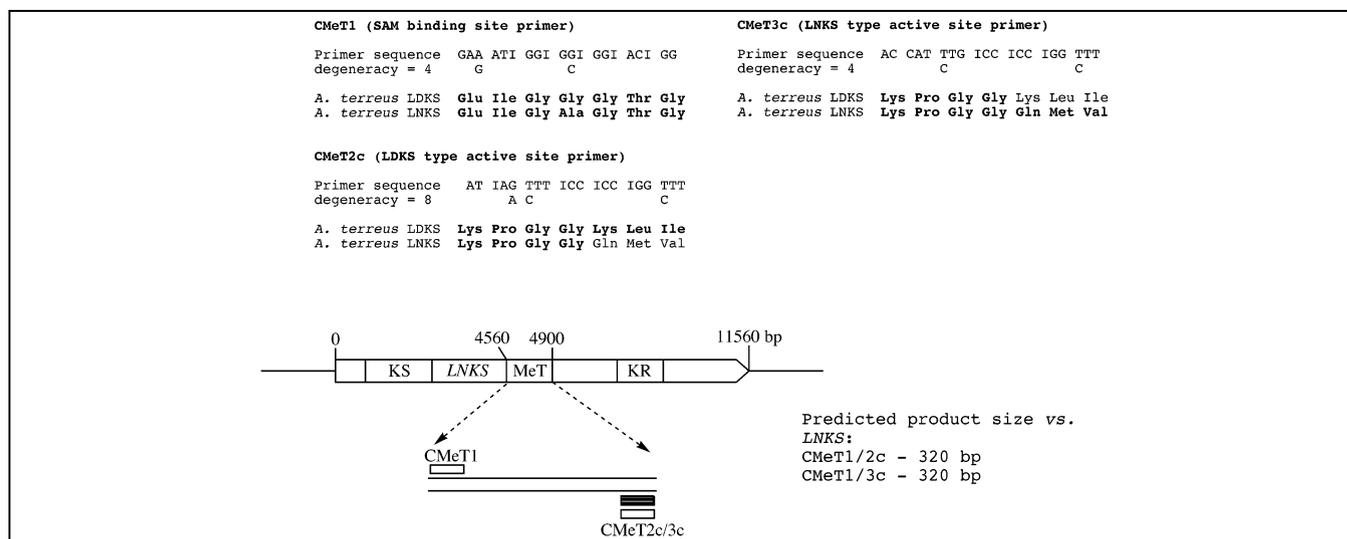


Fig. 8. Design and relative binding positions of CMeT specific primers. Primer sequence and design: Degenerate primer sequence and amino acid sequence of fungal PKS subclass upon which they are based. I = inosine, c = reverse complement nucleotide sequence. Bold text indicates bias to particular sequence conservation motif. Accession numbers for sequences are shown in Table 1.

representing the SAM binding (Motif I) and active/magnesium binding sites (Motif III) (Fig. 8). The nucleotide sequence of the intervening region does not contain introns in either of the two sequences. The following degenerate primers were designed on the basis of these conserved regions: CMeT1, a general SAM binding site primer; CMeT2c, a LDKS biased active site primer; CMeT3c, a LNKS biased active site primer.

Control PCR reactions were run with *A. terreus* genomic DNA (CMeT1/CMeT2c) and pTPKS-100 (CMeT1/CMeT3c) plasmid DNA templates. Both showed multiple products upon gel electrophoresis, under the low annealing temperature PCR conditions used (50°C), including bands in the predicted 300–350 bp size range (see Fig. 9A for a representative example). Bands in the predicted CMeT size range were cloned and sequenced. The CMeT1/

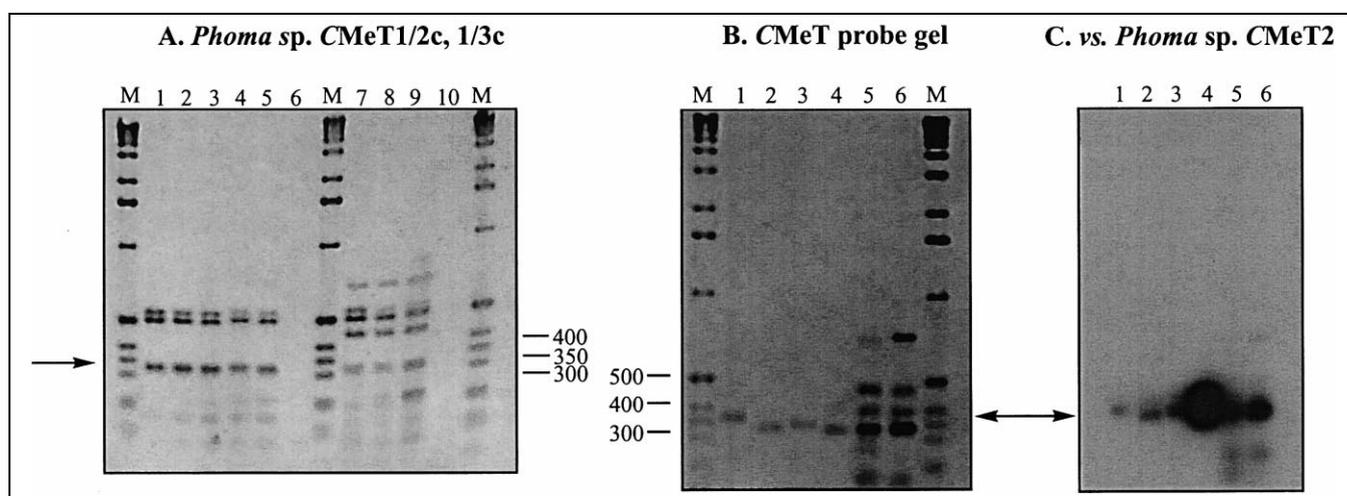


Fig. 9. Detection of PKS CMeT fragments. (A) CMeT1/2c and CMeT1/3c primer PCR with *Phoma* sp. C2932 genomic DNA template. PCR products of 300–400 bp in size, indicated by arrows, were obtained with *Phoma* sp. C2932. Thermal cycling parameters were: (98°C, 5') $\times$ 1, (94°C, 1'; 50°C, 1'; 72°C, 30') $\times$ 34, (72°C, 10'). PCR primers were used at the following concentrations: CMeT1, 1.0  $\mu$ M. CMeT2c, lane 1, 0.5  $\mu$ M; 2, 6.0; 3, 2.0  $\mu$ M; 4, 3.0  $\mu$ M; 5, 4.0  $\mu$ M. CMeT3c, lane 7, 0.5  $\mu$ M; 8, 10, 1.0  $\mu$ M; 9, 2.0  $\mu$ M. Magnesium ion ( $Mg^{2+}$ ) concentration and approximate mass of template DNA used as follows:  $Mg^{2+}$ , 5 mM; genomic DNA, 5 ng (lanes 6/10, 0 ng). Reaction products were fractionated on a 1.5% agarose gel in TAE buffer at 100 mA, 200 V. M, 1 kb DNA ladder size markers (Gibco BRL). (B) Cloned fungal CMeT products obtained by PCR with CMeT1/2c and CMeT1/3c primer pairs from several filamentous fungi were fractionated on a 1.5% agarose gel in TAE buffer at 100 mA/200 V. Samples were analysed as follows: 1, MF5453 CMeT1; 2, MF5453 CMeT2; 3, *Phoma* sp. CMeT1; 4, *Phoma* sp. CMeT2; 5, *P. citrinum* CMeT1/2c PCR; 6, *P. citrinum* CMeT1/3c PCR. M, 1 kb DNA ladder (Gibco BRL). (C) PCR products were blotted from duplicate gels onto nylon membranes and probed with  $^{32}P$ -labelled *Phoma* sp. CMeT1. All hybridisations run and washed at 50°C. Hybridised bands indicated by arrows.





Table 3  
PCR products derived from genomic DNA

|                    | Organism                 | Product/gene     |
|--------------------|--------------------------|------------------|
| KS                 | MF5453                   | LC1/2c (NR)      |
|                    | <i>P. patulum</i>        | LC 1/2c (NR)     |
|                    | MF5453                   | LC3/5c (PR)      |
|                    | <i>P. patulum</i>        | LC3/5c (PR)      |
|                    | <i>Phoma</i> sp.         | pks3 KS3/4c (HR) |
|                    | MF5453                   | pks3 KS3/4c (HR) |
|                    | <i>A. terreus</i>        | pks3 KS3/4c (HR) |
|                    | <i>A. terreus</i>        | LNKS KS3/4c (HR) |
|                    | <i>C. heterostrophus</i> | pks1 KS3/4c (HR) |
|                    | <i>P. citrinum</i>       | KS3/4c PCR       |
|                    | <i>P. patulum</i>        | FAS2 KS          |
|                    | <i>S. coelicolor</i>     | actI             |
|                    | <i>S. erythrae</i>       | eryA ORF1 mod1   |
| KR                 | MF5453                   | KR1/2c (PR)      |
|                    | <i>P. patulum</i>        | KR1/2c (PR)      |
|                    | <i>A. terreus</i>        | LNKS KR7/4c (HR) |
|                    | <i>C. heterostrophus</i> | pks1 KR5/4c (HR) |
|                    | <i>S. coelicolor</i>     | actIII           |
| <i>S. erythrae</i> | eryA ORF1 mod1           |                  |
| CMeT               | MF5453                   | CMeT1, CMET2     |
|                    | <i>Phoma</i> sp.         | CMeT1, CMET2     |
|                    | <i>P. citrinum</i>       | CMeT1/2c PCR     |
|                    | <i>P. citrinum</i>       | CMeT1/3c PCR     |

mology to any known sequences. MF5453 and *Phoma* sp. KR1/KR2c (fungal PR) derived sequences show high homology to each other at both DNA and deduced protein levels (94.0% nt, 88.1% amino acid similarity, 87.5% amino acid identity), suggesting that the genes from which they were derived are functional homologues.

## 2.6. Southern hybridisation analysis

In previous work using PCR primers designed using PR and NR fungal PKS KS regions (LC1/LC2c and LC3/LC5c) we obtained products from PCR reactions using template genomic DNA derived from *P. patulum* and *A. parasiticus* [27]. These PKS gene fragments were sequenced and shown to group with other NR (LC1/LC2c-derived products) and PR (LC3/LC5-derived products) PKS sequences. The PKS fragments were also used in Southern blotting experiments with *Bam*HI- and *Eco*RI-digested genomic DNA from a range of fungal species. The analysis showed that the PCR products could be used as probes in Southern blotting experiments with these heterologous fungi, but that their specificity was such that neither probe class could be used to hybridise to PKS genes from the other class. In fact, it appeared that the PR class of probes was specific for MSAS (i.e. PR) genes, while the NR probes tended to hybridise to spore-pigment (i.e. NR) associated PKS genes. Our analysis also indicated the wide distribution of such PKS genes in filamentous fungi. Although these results showed a clear structural distinction between PR and NR PKS genes it was not apparent whether probes generated using HR KS primers, PR KR

or CMeT primers would display the same behaviour. We therefore set out to examine the utility of our PCR-generated probes in Southern blotting experiments.

### 2.6.1. Structural analysis of probes

The structural relationship between the KS PCR products and gene fragments listed in Table 3 was analysed by probing with PCR products from the three different fungal PKS subclasses represented. These products were agarose-gel-fractionated (Fig. 14A), blotted onto duplicate nylon membranes and probed with <sup>32</sup>P-labelled MF5453 LC1/2c (NR, Fig. 14B), MF5453 LC3/5c (PR, Fig. 14C), MF5453 pks3 KS3/4c (HR, Fig. 14D) and *Phoma* sp. pks3 KS3/4c (HR, Fig. 14E) cloned PCR products. These results show that in general products from one subclass hybridise well to one another and not to products from the other subclasses. In all cases any observed overlap between subclasses could be removed by increasing the post-hybridisation wash temperature used.

The MF5453 NR KS probe hybridises strongly to both its homologous blotted product (Fig. 14B, lane 1) and the equivalent NR KS product obtained from *P. patulum* (Fig. 14B, lane 2). Likewise the MF5453 PR KS probe hybridises strongly to its homologous blotted product (Fig. 14C, lane 3) and to its *P. patulum*-derived equivalent (Fig. 14C, lane 4). Hybridisations with the two HR KS probes, MF5453 pks3 (Fig. 14D) and *Phoma* sp. pks3 (Fig. 14E), also show hybridisation predominantly to HR KS blotted products (Fig. 14D,E, lanes 5–10). Within the HR KS subclass variations in hybridisation intensity can be seen. These can be attributed to the differences in homology between the variety of products obtained using the KS3/4c primer pair. For example the highly homologous *Phoma* sp. pks3 (lane 5) and *A. terreus* LNKS KS (lane 8) products bind more strongly to the *Phoma* sp. KS3/4c probe (Fig. 14E) than to the MF5453 KS3/4c probe (Fig. 14D).

No cross-hybridisation is observed between any of the PKS KS probes and the fungal FAS KS domain (Fig. 14A–E, lane 11). This confirms that sequence homology between the two groups is sufficiently low for them to be structurally distinct and the FAS genes can be discounted as the source of any unknown bands in subsequent genomic hybridisations. The bacterial PKS *act* and *ery* probes (Fig. 14A–E, lanes 12 and 13 respectively) show some cross-hybridisation with the fungal PKS KS probes, suggesting that both Type II and Type I modular PKS probes may have a role to play in fungal PKS gene access.

The structural relationship between CMeT PCR products from Table 3 was determined using *Phoma* sp. CMeT2 radiolabelled PCR product (Fig. 9C). Very strong hybridisation was observed for the *Phoma* sp. CMeT2 probe (Fig. 9C) with all blotted products, from both CMeT subgroups (group 1: MF5453 (lane 1); *Phoma* sp. (lane 3). Group 2: MF5453 (lane 2); *Phoma* sp. (lane 4)) and *P. citrinum* CMeT PCR reaction products (CMeT1/2c

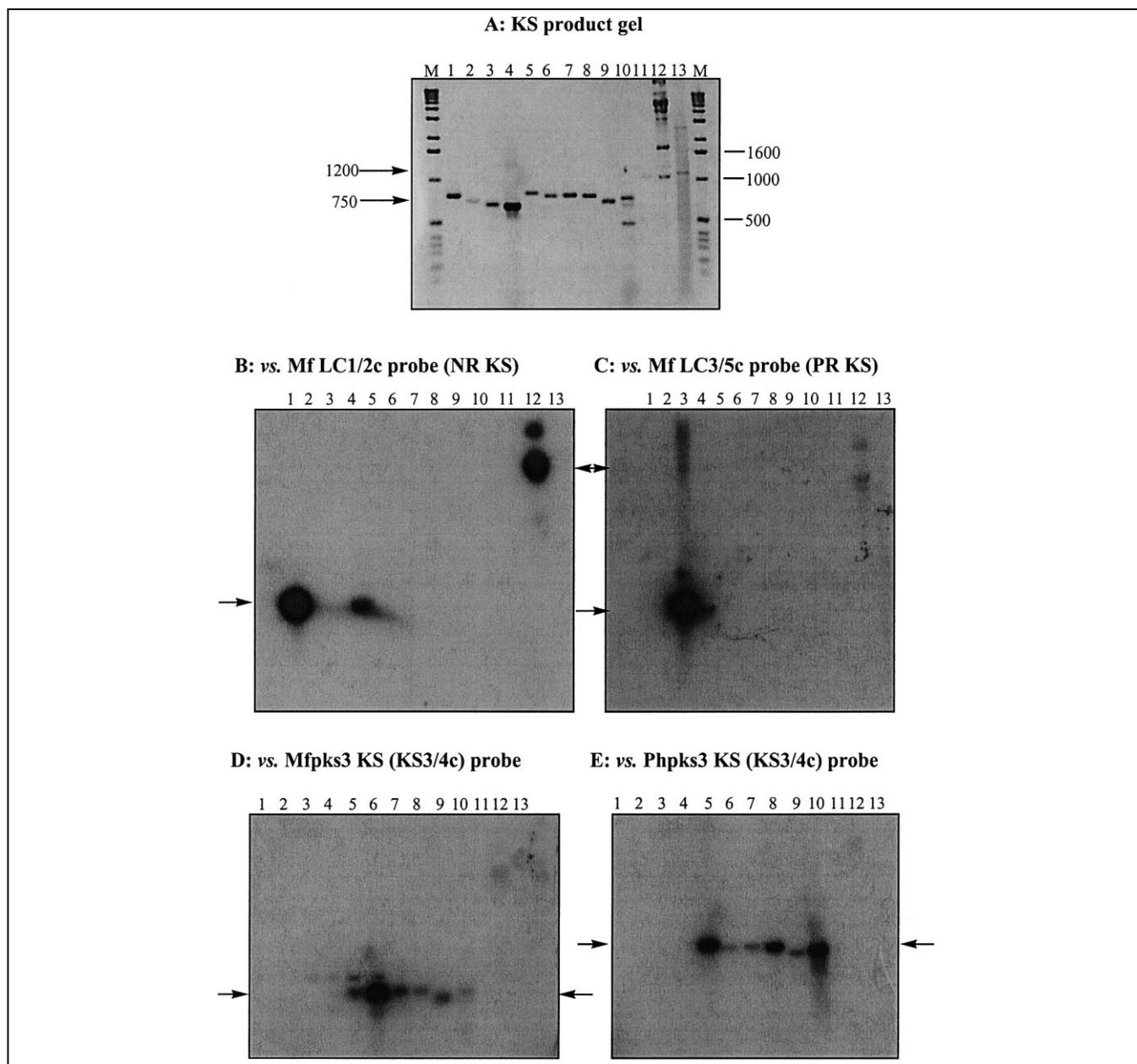


Fig. 14. Southern blot analysis of fungal and bacterial PKS and fungal FAS KS gene fragments with fungal NR, PR and HR PKS PCR primer derived probes. (A) Cloned fungal KS products obtained by PCR with NR (LC1/2c), PR (LC3/5c) and HR (KS3/4c) primer pairs from several filamentous fungi were fractionated along with bacterial PKS and fungal FAS KSs on a 1.5% agarose gel in TAE buffer at 100 mA/200 V. Samples were analysed as follows: 1, MF5453 LC1/2c; 2, *P. patulum* LC1/2c; 3, MF5453 LC3/5c; 4, *P. patulum* LC3/5c; 5, *Phoma* sp. pks3 KS (KS3/4c); 6, MF5453 pks3 KS (KS3/4c); 7, *A. terreus* HR pks3 KS (KS3/4c); 8, *A. terreus* LNKS KS3/4c; 9, *C. heterostrophus* pks1 KS3/4c; 10, *P. citrinum* KS3/4c PCR (REF); 11, *P. patulum* FAS2 KS; 12, *act1*; 13, *eryA* ORF1 mod1 KS. M, 1 kb DNA ladder (Gibco BRL). B–E. PCR products shown in A were blotted from duplicate gels onto nylon membranes and probed with the following  $^{32}$ P-labelled PCR products: B, MF5453 LC1/2c; C, MF5453 LC3/5c; D, MF5453 pks3 KS (KS3/4c); E, *Phoma* sp. pks3 KS (KS3/4c). All hybridisations run and washed at 50°C. Hybridised bands indicated by arrows.

– lane 5, CMeT1/3c – lane 6). The most intense band (lane 4) was seen for the homologous case, but there is clearly sufficient hybridisation to identify the heterologous gene fragments. A similar pattern was observed when probing the products derived from the *P. citrinum* genomic DNA (data not shown).

The structural relationship of KR PCR products and gene fragments listed in Table 3 was also analysed, using

radiolabelled MF5453 KR1/2c (Fig. 15). The MF5453 PR KR probe (Fig. 15B) hybridises strongly to its homologous blotted product (Fig. 15B, lane 1) and the equivalent PR KR product obtained from *P. patulum* (Fig. 15B, lane 2). Cross-hybridisation with the much less homologous (by sequence analysis) bacterial *actIII* KR gene is also evident, as well as weak hybridisation to the *ery* clone. This presumably indicates some underlying DNA sequence homol-

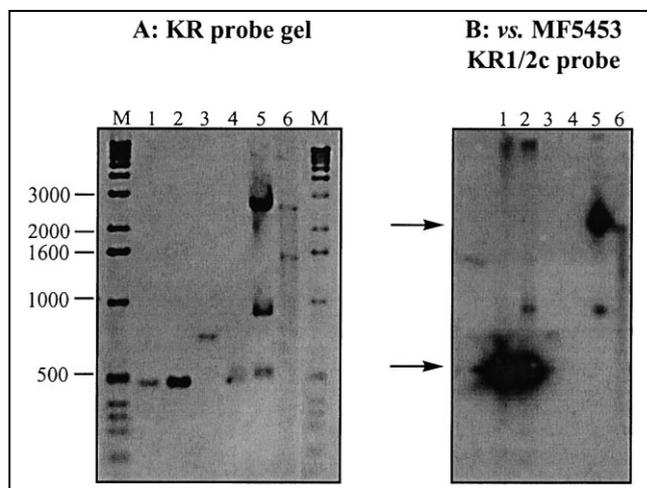


Fig. 15. Southern blot analysis of fungal and bacterial PKS KR gene fragments with fungal PR and HR PKS PCR primer derived probes. (A) Cloned fungal KR products obtained by PCR with PR (KR1/2c) primer pairs from several filamentous fungi were fractionated along with bacterial PKS KR on a 1.5% agarose gel in TAE buffer at 100 mA/200 V. Samples were analysed as follows (described in this report unless otherwise stated): 1, MF5453 KR1/2c; 2, *P. patulum* KR1/2c; 3, *A. terreus* LNKS KR7/4c; 4, *C. heterostrophus* pks1 KR5/4c; 5, *actIII*; 6, *eryA* ORF1 mod1 KR. M, 1 kb DNA ladder (Gibco BRL). (B) PCR products shown in (A) were blotted from duplicate gels onto nylon membranes and probed with the  $^{32}\text{P}$ -labelled MF5453 KR1/2c. PCR products shown in lanes 3 and 4 were derived from failed attempts to design and use HR KR primer pairs. All hybridisations run and washed at 50°C. Hybridised bands indicated by arrows.

ogy not predicted by our analysis – or it may simply indicate very high homology across a short sequence, such as that encoding the NADPH binding site.

### 2.6.2. Southern blotting of genomic digests

Genomic DNA from MF5453 and *Phoma* sp. was digested with *Bam*HI and *Eco*RI and duplicate versions blotted onto nylon membranes (Fig. 16). Hybridisation of these genomic Southern blots was performed using the  $^{32}\text{P}$ -labelled PCR products described for the PCR product blotting.

Hybridisations with different PKS domain probes show a number of key features. All KS-probed digests (Fig. 17A–D) contain at least one hybridised restriction fragment with each probe used except for the *Phoma* sp. digest vs Mfpks3 KS probe (Fig. 17C, lanes 3 and 4). The Mfpks1 NR KS probe shows at least two hybridised fragments in both the MF5453 (Fig. 17A, lanes 1 and 2) and *Phoma* sp. (Fig. 17A, lanes 3 and 4) genomic blots. A single homologous band was observed in both the MF5453 (Fig. 17B, lanes 1 and 2) and *Phoma* sp. (Fig. 17B, lanes 3 and 4) genomic digests when probed with the Mfpks2 PR KS probe.

Probing the genomic Southern blots with HR KS probes gave rise to different hybridisation patterns for the Mfpks3 (Fig. 17C) and Phpks3 (Fig. 17D) radiolabelled products. For the Mfpks3 probe a single, presum-

ably homologous, hybridised restriction fragment was observed with the MF5453 digest (Fig. 17C, lanes 1 and 2), whereas no bands were observed with the *Phoma* sp. digest (Fig. 17C, lanes 3 and 4). The absence of any bands from the *Phoma* sp. digest suggests that the *Phoma* sp. genome does not contain any PKS genes of sufficient homology to the Mfpks3 product for hybridisation to occur. If the Mfpks3 KS gene fragment is derived from one of the putative SQS genes then the presumed similarity between these genes and those encoding SQS in *Phoma* sp. might have suggested that at least one gene of high homology to the Mfpks3 KS fragment would be present on its genome. The lack of hybridised restriction fragments observed cannot however rule out the possibility of the Mfpks3 KS fragment originating from a SQS PKS gene.

Probing with the Phpks3 product revealed a single hybridised band for the digests of both MF5453 (Fig. 17D, lanes 1 and 2) and *Phoma* sp. (Fig. 17D, lanes 3 and 4). This observation was not unexpected, despite the high apparent specificity of HR KS probes, as the MF5453 genome had been shown by PCR to contain a gene fragment (Mfpks4) showing high homology to the Phpks3 KS gene fragment. It should be noted throughout the KS probe genomic blotting that no overlap is observed between the bands obtained with different subclass probes. This indicates that the probe class specificity demonstrated by structural analysis has been maintained.

CMeT probe genomic blot hybridisations show multiple high intensity bands with the *Phoma* sp. CMeT1 (Fig.

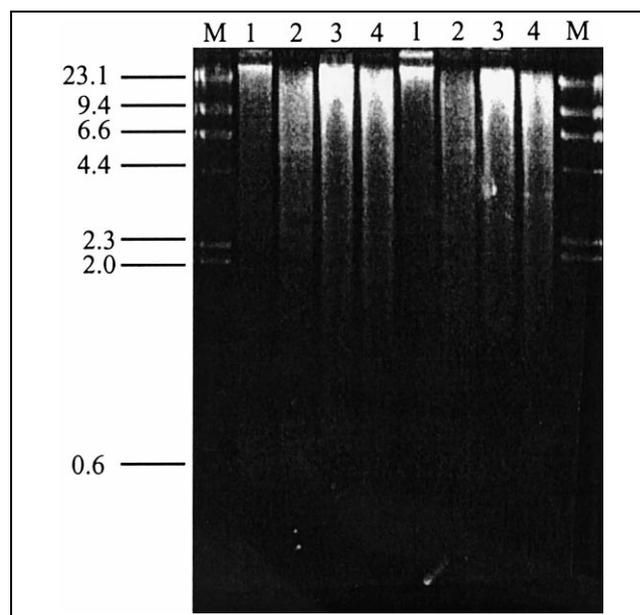


Fig. 16. Restriction digests of MF5453 and *Phoma* sp. genomic DNA for Southern blotting. Genomic DNA from MF5453 (1, 2) and *Phoma* sp. C2932 (3, 4) was digested overnight at 37°C with restriction enzymes *Bam*HI (1, 3) and *Eco*RI (2, 4). Approximately 5 µg of DNA from each digest was fractionated on a 1% agarose gel in TAE buffer at 30 mA/30 V. Each sample was applied to the gel in two separate lanes. M,  $\lambda$  HindIII DNA size markers.

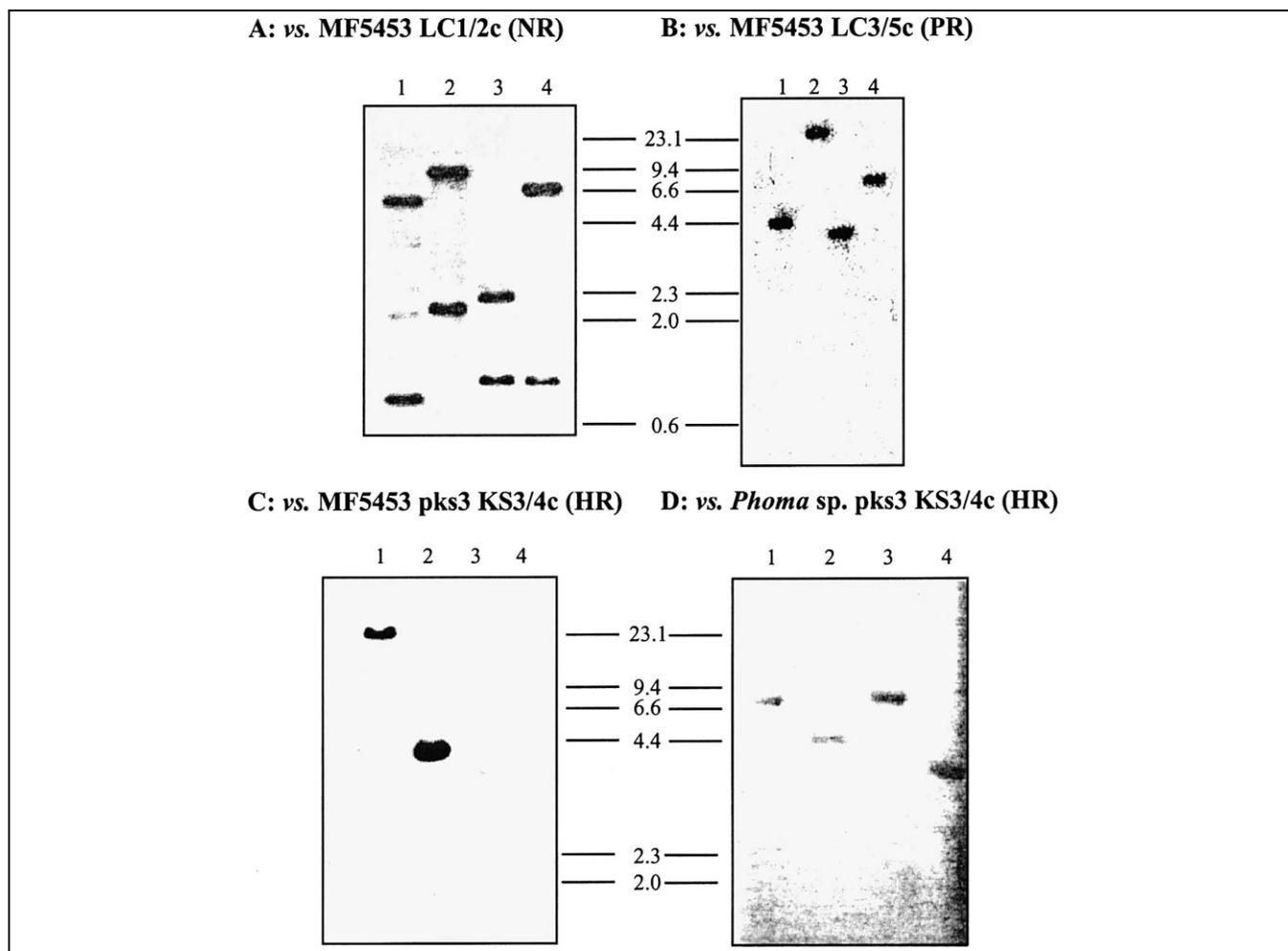


Fig. 17. MF5453 and *Phoma* sp. genomic Southern blots probed with fungal PKS KS PCR products. The MF5453 and *Phoma* sp. restriction digests shown in Fig. 16 were blotted onto duplicate nylon membranes and probed with the following  $^{32}\text{P}$ -labelled KS PCR products: (A) MF5453 LC1/2c (NR); (B) MF5453 LC3/5c (PR); (C) MF5453 pks3 KS (KS3/4c); (D) *Phoma* sp. pks3 KS (KS3/4c). Probes were hybridised and washed at 55°C. The relative positions of the  $\lambda$  HindIII DNA size markers on the hybridised membrane are indicated between the panels (in kb).

18A). The probe derived from MF5453 gave identical results (data not shown). The main hybridisation bands for the MF5453 (*Bam*HI: 4.6, 6.2, 7.0 and 8.2 kb (lane 1). *Eco*RI: 3.4, 4.4 and 7.0 kb (lane 2)); and *Phoma* sp. (*Bam*-HI: 4.6, 6.0, 9.0 kb (lane 3). *Eco*RI: 3.5, 4.4, 8.8 kb (lane 4)) genomic digests are observed with both CMeT probes. This suggests that CMeT probes are capable of binding to restriction fragments containing a variety of CMeT-like genes and are hence less specific than the KS and KR probes. It is unclear if any overlap between these bands and those observed with the HR KS probes (Fig. 17C,D) exists, due to the less well defined position of the CMeT bands.

KR probe hybridisations show hybridised restriction fragments with the MF5453 PR KR probe (Fig. 18B). The presence of single strong bands in the MF5453 genomic digests (Fig. 18B, lanes 1 and 2) and also single bands in the *Phoma* sp. genomic digests (Fig. 18B, lanes 3 and 4) was to be expected as the PR KR PCR products

obtained previously by PCR from both organisms were highly homologous. The good homology observed between the MF5453 PR KR product and known *MSAS* genes would appear to indicate the presence of a single *MSAS*-like gene on both genomes. Interestingly these bands appear to overlap with those observed for the equivalent PR KS probe hybridisations (Fig. 17B). This may indicate that the KS and KR *MSAS*-like gene fragments are present on the same PKS gene on the MF5453 and *Phoma* sp. genomes.

### 3. Discussion

Numerous strategies have been used to clone PKS genes from their hosts, but none so far offers a general and selective method for obtaining PKS genes from widely varying species of fungi. Recently a new and innovative method involving shotgun sequencing of genomic DNA

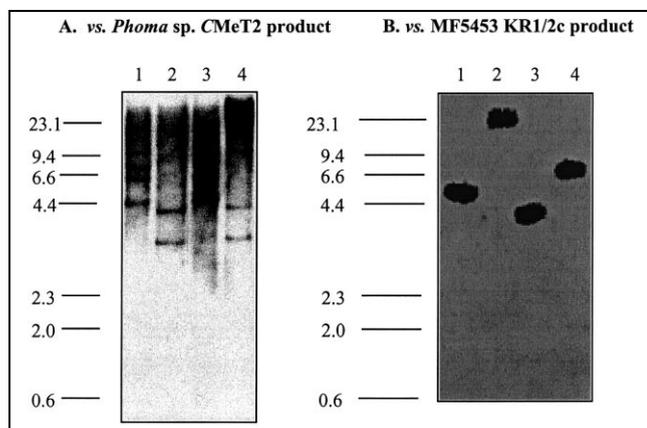


Fig. 18. (A) MF5453 and *Phoma* sp. genomic Southern blots probed with *Phoma* sp. PKS CMeT2 PCR products. The MF5453 and *Phoma* sp. restriction digests shown in Fig. 16 were blotted onto duplicate nylon membranes and probed with  $^{32}\text{P}$ -labelled *Phoma* sp. CMeT1; Probes were hybridised and washed at 55°C. The relative positions of the  $\lambda$  *Hind*III DNA genomic digest gel size markers on the hybridised membrane are indicated between the panels (in kb). (B) MF5453 and *Phoma* sp. genomic Southern blots probed with MF5453 KR1/2c PCR product. The MF5453 and *Phoma* sp. restriction digests shown in Fig. 16 were blotted onto duplicate nylon membranes and probed with  $^{32}\text{P}$ -labelled MF5453 KR1/2c (PR). The probe was hybridised and washed at 55°C. The relative positions of the  $\lambda$  *Hind*III DNA size markers on the hybridised membrane are indicated to the side of the panel (in kb).

has been shown to be effective for cloning modular PKS from myxobacteria, but this is effective only when the length of the PKS genes is significant [34]. The PKS genes of interest must compose 0.5–1% of the genome to make the method viable. This is simpler when the genome is relatively small (ca 10 Mb for *S. cellulosum*) and when the PKS is relatively large (50 kb for the epothilone cluster), as is the case with modular PKS in actinomycete hosts. This is less of a practical solution for fungal hosts containing relatively small Type I PKS genes – most Type I PKS are in the 6–8 kb range with a ‘typical’ genome from a filamentous fungi being 25.4 Mb for *A. nidulans* [35].

The use of heterologous gene probes could also be envisioned, but we and others have shown that the commonly used bacterial Type II and modular PKS probes are ineffective in fungi. The recently cloned fumonisins B1 synthase from *G. fujikuroi* was isolated using a cloning procedure relying on a homologous probe generated from cDNA obtained from a producing organism [26]. However the PCR primers used to generate the probe were not biased towards a particular class of PKS (in fact they were general KS primers) and these workers were fortunate not to have isolated fragments from the ubiquitous THNS and MSAS genes.

We have sought to design gene probes which could overcome some of the problems discussed above. In previous work we showed that fungal PKS genes could be divided into classes based on the level of reduction of their encoded polyketide. With the growth in the number of

known fungal PKS sequences we have been able to use sequence data to design new probes to expand the range of PKS types. Using a relatively straight forward approach we manually identified short regions of high homology which held within classes, but were selective enough to not specify sequences from outside the class. This process was repeated at two positions on any given sequence so that the cognate DNA sequences could be used as primers for PCR reactions. We used degenerate DNA primers containing mixed bases or inosine to maximise the chances of successful PCR from unknown PKS genes, and to allow for unknown codon preferences.

In this way we designed three different sets of primers specifying three different classes of fungal PKS KS regions: NR, such as orsellinic acid **10**, THN **3**, YwA1 **5** etc; PR, such as 6MSA; and HR such as T-toxin **6** and lovastatin **7a**. We used these on isolated genomic DNA from control organisms known to contain PKS genes of high homology to the primers and also on strains known to produce polyketides but where there were no known PKS gene sequences. Thus when using NR selective primers we amplified gene fragments very highly homologous to known THNS genes from *Phoma* sp. and MF5453. Using PR selective primers we amplified fragments of putative MSAS genes from these organisms also. These results were somewhat expected as we have already begun to discover that THNS and MSAS appear to be ubiquitous in filamentous fungi.

We also designed primers selective for the HR class of fungal PKS. These HR PCR primers worked well with the control strain *A. terreus* where we amplified a fragment of the known LNKS. We also amplified a putative fragment from the *P. citrinum* compactin synthase. More importantly we discovered new gene fragments from putative PKS genes when using HR specific primers with genomic DNA from *Phoma* sp. and MF5453. These gene fragments were homologous with other known HR PKS genes. The TreePlot (Fig. 7) clearly shows how the NR, PR, HR, and modular PKS genes group together.

In Southern blotting experiments we have shown that the KS products derived from the PCR experiments cross-hybridise only with other members of their class, e.g. the HR PCR product from MF5453 genomic DNA hybridises strongly with the HR fragments from MF5453, *Phoma* sp., *A. terreus*, LNKS plasmid, *C. heterostrophus* and *P. citrinum* (Fig. 14D). Interestingly other hybridised bands were evident that were not clear in the electrophoresis gel of the products (Fig. 14A). For example PCR products from *Phoma* species (Fig. 14A, lane 5) showed one clear product band on the gel, but two hybridised bands when probed with the radiolabelled MfPKS3 PCR product (Fig. 14D, lane 5). This perhaps indicates that not all products from a genomic PCR are amplified to the same extent and our PCR products are probably heterologous, containing more than one sequence in some cases.

This experiment was an important control for the ge-

nomic Southern blots (Fig. 17). In the genomic blots we observed multiple hybridised bands for NR selective probes in both MF5453 and *Phoma* species. The presence of multiple bands can be attributed to the presence of additional NR subclass (spore pigment/melanin) genes. Multiple PKS-derived pigments are known to be produced by individual filamentous fungi, for example Bingle et al. amplified two distinct NR KS gene fragments from *A. parasiticus* using the LC1/2c (NR) primer pair [27] and Hohn and coworkers isolated seven PKS fragments (none of which were the desired FUMS) from genomic DNA of *G. fujikuroi* [26]. For PR selective probes we found only a single hybridised band in each species; this is presumably the endogenous MSAS from each of these organisms.

The HR KS probes were also selective in genomic Southern blots. The HR KS PKS gene fragment derived from MF5453 using the HR selective PCR primers hybridised well with its homologous genomic DNA, but under our standard hybridisation conditions it did not hybridise well with *Phoma* genomic DNA. The *Phoma*-derived probe was more useful, however, giving single hybridised bands with genomic DNA from both species.

In comparing the Southern blots of the genomic DNA it appears that none of the hybridised bands is repeated in a different probing experiment. This suggests that for the MF5453 species there are at least five PKS genes on the genome (2×NR, 1×PR and 2×HR), while the *Phoma* species contains at least four different PKS genes (2×NR, 1×PR and 1×HR). This compares well with the number of isolated PCR products for each species (i.e. four for MF5453 and three for *Phoma* sp.) but is still likely to be an underestimate for the *Phoma* species as we would expect two HR PKS to be involved during the biosynthesis of squalestatin S1 **9**. In the case of *A. terreus* there must be at least five PKS genes (two from LNKS/LDKS, two more HR described here and the known MSAS [17]). This compares with the seven known polyketides produced by this fungus: the diketide sidechain from mevinolin **7a**; the tetraketides 6MSA **2** and orsellinic acid **10**; the pentaketide isocoumarins such as dihydrocitrinone **11**; octaketides such as geodin **12**; and the nonaketide of lovastatin **7a**.

We also examined the possibility of developing CMeT specific PCR primers. At the time of our investigation sequences for LNKS and LDKS were the only available CMeT in the database. Despite this, these PCR primers successfully amplified gene fragments highly homologous to CMeT sequences from various fungi. Interestingly *P. citrinum* genomic template DNA gave multiple bands, more than one of which hybridised when probed with radiolabelled CMeT fragment (Fig. 9B,C). *P. citrinum* may not contain a LNKS type CMeT as compactin **7b** lacks methylation on the nonaketide. However we would expect other PKS CMeT genes to be present on the genome, to account for the compactin diketide methyl group

and also the three C-methyl groups of sclerotinin A **13** which is a known metabolite of *P. citrinum*.

The CMeT products from MF5453 and *Phoma* sp. were cloned and sequenced. These sequences are compared in the TreePlot (Fig. 11) along with newly deposited sequences for other CMeT such as those from the mithramycin [36], and epothilone [10] biosynthetic gene clusters. This shows three groups of CMeT sequences: a fungal 'Type I' group represented by LNKS, LDKS, our PCR products and the fumonisin sequence; a 'modular' type group represented by the epothilone synthase and a CMeT domain from a putative modular *Bacillus subtilis* synthase; and discrete post-PKS types such as the mithramycin CMeT and various NMeT and OMeT.

In our experiments we isolated two CMeT fragments from each organism. Sequence comparison indicates that these sequences fall into two groups: LDKS-like and LNKS-like. The genomic blot (Fig. 18A) indicates the presence of at least four hybridising bands in each organism. The newly published CMeT sequences from the epothilone cluster and mithramycin, together with our four new CMeT PCR product sequences now form a significant group and it should be possible to improve the specificity of CMeT primers to perhaps select for individual classes, e.g. 'Type II', 'modular', or 'fungal Type I'.

Parallel experiments attempting to use conserved homology in KR PKS domains were much less successful. We were able to make PR specific primers for use in genomic PCR and we amplified the expected fragments of putative MSAS genes. The fungal PR-biased primers were much more successful than primers designed on the KR domains from actinomycete modular PKS. In genomic blots the PR fungal-primer-derived fragments hybridised well to what appeared to be the same bands visualised when probing with PR KS probes. Presumably the KS and KR gene fragments are derived from the same genomic ORF, and perhaps both the KS and KR are from each organism's endogenous MSAS since both sets of products show high homology to known MSAS sequences from the database. Unfortunately HR KR primers were much less useful: the PCR products from these primers were not significantly similar to known KR sequences and the resulting probes were not useful in Southern blots.

#### 4. Significance

The sequences of PKS selective PCR primers for fungal PKS genes described here should be useful for a number of purposes. Foremost, they should assist any fungal PKS isolation strategy requiring a high measure of selectivity in avoiding the cloning of one of the myriad of 'unwanted' fungal PKS encoded by typical fungal genomes. We are currently examining the use of HR KS and CMeT probes in the cloning of the SQS genes from *Phoma* sp. C2932 and MF5453. Another use could be the screening of commercial

fungal libraries to determine their potential for mycotoxin or other bioactive compound production. We are currently examining the use of these PCR primers for the generation of probes from cDNA synthesised at different time-points during the fungal growth cycle to determine the temporal relationships between production of known polyketides and PKS gene expression. The use of selectively generated homologous probes for KS domains, in conjunction with our fungal PKS specific CMET probes should increase the likelihood of cloning any given fungal PKS. Our strategy for the design and production of gene probes for secondary metabolite production could be extended to other important classes of secondary metabolite synthases, such as the non-ribosomal polypeptide synthases (NRPS) where many sequences are known and a high level of understanding has been achieved in linking protein sequence to substrate usage and potential product structure [37].

## 5. Materials and methods

Standard cloning procedures were adopted and manufacturer's instructions followed unless otherwise indicated [28]. Media constituents were purchased from Difco unless otherwise stated. Milli Q purified water was used for all media and solutions. All chemicals, reagents and buffer ingredients used were of ANALAR or molecular biology grade and were purchased from either Sigma Chemical Co. Ltd. or BDH Chemicals Ltd. Ultrapure agarose was obtained from Gibco BRL. Molecular biological kit reagents were used as supplied by the manufacturer. Restriction enzymes, T4 DNA ligase, Alkaline Phosphatase and other modifying enzymes were obtained from Boehringer Mannheim. *Taq* DNA polymerase and T4 polynucleotide kinase were obtained from Promega. Bacterial cultures were grown overnight at 37°C, unless otherwise stated. Liquid cultures were grown in a rotary incubator at 250 rpm. Fungal mycelial cultures were produced by inoculation of liquid medium from cryogenically frozen glycerol seed suspensions or paraffin-stored slopes. Culture conditions were as follows, unless otherwise stated. Seed flasks were inoculated and incubated at 25°C and 250 rpm for 3 days. Aliquots (1 ml) of the seed culture were transferred to fresh medium and grown for a further 3–5 days under the same conditions.

### 5.1. Unidentified fungus (*Merck strain MF5453*)

Tomato paste (J. Sainsbury) (40.0 g), corn steep liquor (5.0 g), D-glucose (10.0 g), oatmeal (J. Sainsbury) (10.0 g) and trace mineral solution ( $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ , 1.0 g l<sup>-1</sup>;  $\text{MnSO}_4 \cdot 4\text{H}_2\text{O}$ , 0.758 g l<sup>-1</sup>;  $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ , 0.025 g l<sup>-1</sup>;  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ , 0.1 g l<sup>-1</sup>;  $\text{H}_3\text{BO}_3$ , 0.056 g l<sup>-1</sup>;  $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24} \cdot 4\text{H}_2\text{O}$ , 0.019 g l<sup>-1</sup>;  $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ , 0.2 g l<sup>-1</sup>) (10 ml), adjusted to pH 6.8 (NaOH, 2 M). Aliquots of 50 ml were taken and sterilised by autoclave.

### 5.2. *Phoma* sp

SC4/SC5 media: Lactose (100 g l<sup>-1</sup>), cotton seed flour (20 g

l<sup>-1</sup>) and 0.5% Junlon solution (20 ml l<sup>-1</sup>, SC4 only) was heated to just below boiling point. Aliquots of 50 ml were taken and sterilised by autoclave in 250 ml conical flasks. SC5 medium was inoculated from SC4 after 3 days.

### 5.3. *P. patulum*

Potato dextrose broth (Difco) was made up according to the manufacturer's instructions and aliquots (50 ml) were sterilised by autoclaving in 250 ml conical flasks. Cultures were incubated at 27–28°C, 200 rpm for 5 days.

### 5.4. *A. terreus/C. heterostrophus*

Fungal seed (FS) medium: peptone (Oxoid) (10 g l<sup>-1</sup>), malt extract (21 g l<sup>-1</sup>) and glycerol (40 g l<sup>-1</sup>). Aliquots of 50 ml were placed in 250 ml conical flasks, to which several 1 mm glass beads (BDH) were added prior to autoclaving.

### 5.5. *P. citrinum*

Seed medium: corn steep liquor (12.5 g l<sup>-1</sup>), glucose (50.0 g l<sup>-1</sup>), calcium carbonate (2.0 g l<sup>-1</sup>) and soy bean oil (2.5 g l<sup>-1</sup>), pH adjusted to 6.3 and aliquots autoclaved separately (50 ml per 250 ml flask). Seed cultures were incubated at 26°C for 48 h at 240 rpm. Production medium (compactin): corn steep liquor (10 g l<sup>-1</sup>), glucose (60 g l<sup>-1</sup>) and soy bean oil (1.0 g l<sup>-1</sup>), pH 6.3, autoclaved in 50 ml aliquots in 250 ml conical flasks. Production medium inoculated with 5 ml seed culture, production starts after 48 h and peaks after 72 h.

Mycelia were separated from the extracellular fluid by centrifugation (Sorvall Instruments RC5C centrifuge, rotor GS-3) at 6000 rpm, 4°C, for 15 min. The mycelia were filtered under vacuum, dried, flash frozen in N<sub>2</sub>(l) and stored at -20°C. If required, seed suspensions were stored under liquid nitrogen by taking 1.0 ml aliquots of the seed culture, mixing with 1.0 ml 40% glycerol and flash freezing in 2.0 ml cryovials (Corning). Slopes were preserved at -20°C under liquid paraffin.

### 5.6. Fungal genomic DNA extraction

Mycelium (approximately 4.0 g) was ground, with pure sand, to a fine powder under N<sub>2</sub>(l) using a pre-cooled pestle and mortar. The ground mycelium was placed in a 35 ml Nalgene polypropylene Oak Ridge centrifuge tube and resuspended at 0°C in 10 ml of DNA extraction buffer (10 mM Tris buffer pH 8.0, 10 mM EDTA, 0.5% SDS). Phenol:chloroform:isoamyl alcohol (25:24:1, 10 ml) was added to the aqueous mycelium solution and mixed slowly for 15–30 min on a horizontal cylindrical rotor (Denley Spiramix 5). The phases were separated by centrifugation (Sorvall Instruments, SS-34 rotor) at 6000 rpm for 15 min at 4°C. The aqueous layer was removed and the phenol extraction/aqueous layer separation procedure repeated until the interface between the two layers was clear. Any traces of phenol were removed by treating the aqueous layer with a chloroform:isoamyl alcohol mixture (24:1) and the phases were separated as before.

Ribonuclease A (bovine pancreas) solution (10  $\mu\text{l}$ , 20.0  $\text{g l}^{-1}$ ) was added to the aqueous layer and incubated at 37°C for 30 min. Phenol extraction, followed by chloroform extraction was repeated once more and the DNA precipitated from the aqueous solution with 2.5 volumes of 100% ethanol and 1/10 volume of LiCl solution (4 M) at –20°C overnight. The DNA was recovered by centrifugation at 13 000 rpm for 10 min, the pellet washed with 70% ethanol at 4°C, air-dried and resuspended overnight in approximately 1 ml TE buffer (10 mM Tris–HCl, pH 8.0, 1 mM EDTA). Genomic DNA for use as a template in PCR reactions was stored at 4°C.

Genomic DNA for use in Southern Blotting was further purified by two rounds of equilibrium centrifugation in a continuous CsCl–ethidium bromide gradient (TE buffer containing CsCl (density = 1.56  $\text{kg l}^{-1}$ ) and ethidium bromide (0.4  $\text{g l}^{-1}$ ). Samples were centrifuged at 100 000 rpm for approximately 16 h at 15°C using a TL00 ultracentrifuge (Beckman) with a TLA-100.2 fixed angle rotor. The genomic DNA band was visualised with a UVL-21 long wave lamp (Ultra-Violet Products Inc.), removed with a hypodermic needle and syringe and diluted in 2.5 volumes of water. Ethidium bromide and CsCl were removed by the addition of 0.6–1.0 volumes of ice-cold isopropanol and precipitation at –20°C for 15 min. DNA was pelleted by centrifugation for 10 min at 13 000 rpm. The DNA pellet was dried and resuspended in 200  $\mu\text{l}$  of TE overnight at 4°C.

### 5.7. Oligonucleotide synthesis

Oligonucleotides were synthesised using a Pharmacia Gene Assembler Plus instrument. Oligonucleotides were deprotected and cleaved from the support matrix by overnight incubation at 55°C in 1 ml ammonia. Blocking groups and ammonia were removed by chromatography on a Pharmacia NAP10 column and the oligonucleotide was eluted in 1.5 ml water. The concentration was estimated spectrophotometrically using the factor 1 A.U. = 20  $\mu\text{g ml}^{-1}$ , at 260 nm absorbance.

### 5.8. PCR

Oligonucleotide-specified sections of DNA were routinely amplified by PCR using *Taq* DNA polymerase, giving 3'A overhung products, in a final volume of 50  $\mu\text{l}$  under the following standard conditions. Reaction solutions contained 1 $\times$ PCR buffer (10 mM Tris–HCl, 1.5 mM  $\text{MgCl}_2$ , 50 mM KCl, pH 8.3 (20°C)), primers at 0.5–1.0  $\mu\text{M}$ , dNTPs at 200  $\mu\text{M}$ , 1 U of enzyme and 10–100 ng genomic or 0.05–1 ng plasmid DNA template. Samples were overlaid with mineral oil to prevent evaporation and amplified in a programmable thermal controller (Genetic Research Instruments Ltd., PTE-100). An initial denaturation at 98°C for 3–5 min was performed and the PCR mixtures allowed to cool to 60°C before addition of the deoxynucleotides and polymerase. PCRs were optimised by addition of  $\text{MgCl}_2$  solution (25 mM, 0–7  $\mu\text{l}$ ) at  $\geq 60^\circ\text{C}$ , to give a total  $\text{Mg}^{2+}$  concentration of 1.5–5.0 mM. The following thermal cycling profile was run over 34 cycles, denaturation at 94°C for 1 min, primer annealing at 45–55°C for 1 min and extension at 72°C for 1–5 min, allowing 1 min

per kb of DNA to be amplified. This was followed by a final extension at 72°C for 10 min. Control reactions without primers and without template were also run for each set of PCR primer pair reactions run.

### 5.9. Gel purification of DNA fragments (PCR products)

Size-fractionated DNA fragments were purified from agarose gels by use of either the High Pure PCR Purification or Agarose Gel DNA Extraction kits (Boehringer Mannheim, 400–5000 bp size fragments), or by the gel electroelution method described below (fragments > 5000 bp, large plasmids). The DNA fragment of interest was excised in the minimum possible volume from an agarose gel of appropriate concentration, made and run in 0.5 $\times$ TBE buffer (0.045 M Tris–borate, 0.001 M EDTA). The excised slice was soaked in 0.5 $\times$ TBE for 10 min and electroeluted using a 0.5 $\times$ TBE buffered IBI unidirectional electroelutor utilising a  $\text{NH}_4\text{OAc}$  (7.5 M)/bromophenol blue salt cushion (75  $\mu\text{l}$ ). The electroelutor was run at 90 V, 1 mA, for a time correlated with the size of the DNA fragment being purified. The salt cushion/DNA solution (400  $\mu\text{l}$ ) was removed and DNA precipitated with 2.0 volumes of ethanol at –20°C for at least 2 h. The eluted DNA was centrifuged at 13 000 rpm for 10 min and the pellet, washed, dried and resuspended in TE buffer.

### 5.10. Ligation reactions

Ligations were performed as described below for pBluescript plasmid DNA or according to the manufacturer's instructions using the pGEM-T vector system. Approximately 100 ng pBluescript plasmid DNA was ligated with insert DNA (molar ratio, vector:insert, 1:3) in the presence of 1 $\times$ ligase buffer and 1 unit T4 DNA ligase in a total volume of 10  $\mu\text{l}$ . Single cut 'sticky ended' ligation reaction solutions were warmed to 45°C for 5 min prior to addition of buffer/ligase and chilled to 0°C. Ligations were run overnight at either 16°C (blunt ended) or 4°C (sticky ended). The ligase was inactivated by incubation at 65°C for 10 min and the ligated DNA ethanol-precipitated and resuspended in 10  $\mu\text{l}$   $\text{dH}_2\text{O}$ .

### 5.11. Sequencing

Samples of concentration 100 ng  $\mu\text{l}^{-1}$  were submitted for sequencing with T3/T7 (pBluescript) or M13 uni/rev (pGEM-T) primer pairs. DNA sequencing was performed by Rhiannon Murray, Department of Biochemistry, University of Bristol (Bristol, UK), using *Taq* DNA polymerase cycle sequencing (temperature cycling, dideoxynucleotide termination) with fluorescent dye detection.

### 5.12. Transfer of DNA to nylon membranes

DNA for Southern blotting was transferred to positively charged nylon membrane (Boehringer Mannheim) by alkaline capillary transfer for 5–16 h, using 0.4 M NaOH as the transfer buffer. Gels containing large DNA fragments were soaked in 0.25

M HCl for 15 min and rinsed with water prior to blotting, to allow partial depurination of the DNA.

### 5.13. Preparation of radiolabelled DNA probes

Cloned DNA inserts were excised from recombinant plasmids by restriction digestion. Fragments were isolated by gel purification and purified using the High Pure DNA purification system (Boehringer Mannheim). Probe DNA was labelled with [ $\alpha$ - $^{32}$ P]-dCTP using the Ready To Go<sup>™</sup> DNA-Labeling (-dCTP) reaction mix (Pharmacia). Labelled probe DNA was separated from unincorporated radioisotope using Nick columns (Pharmacia) and eluted in 400  $\mu$ l water. Finally the labelled DNA was denatured by boiling for 5 min, cooled on ice and either used immediately or stored at  $-20^{\circ}\text{C}$  for future use.

### 5.14. Hybridisation of radiolabelled probes to Southern blots

The blotted membrane was prehybridised in 40 ml hybridisation buffer (0.25 M phosphate buffer pH 7.2, 7% SDS) for 5–15 min at the appropriate temperature. This buffer was replaced with 10 ml of fresh buffer and the probe DNA (200  $\mu$ l) added. Hybridisation reactions were incubated overnight in an HB-1 hybridising oven (Techne). After discarding the hybridisation buffer the membrane was rinsed briefly in wash buffer ( $2\times\text{SSC}$  (150 mM NaCl, 15 mM sodium citrate, pH 7.0), 0.1% SDS) followed by washing for 15 min in the same buffer. Unless otherwise stated hybridisation and washing steps were performed at  $55^{\circ}\text{C}$ . The membrane was dried, wrapped in Saran<sup>™</sup> wrap (Dow) and fluorescent paint positional markers attached next to the membrane to allow easy orientation of the resulting autoradiograph. After post-hybridisation washing, membranes were exposed to hyperfilm MP film (Amersham) in a cassette with an intensifying screen at  $-70^{\circ}\text{C}$ . After an appropriate exposure time autoradiographs were developed using LX-24 X-ray developer (Kodak) and Hypam rapid fixer (Ilford). Probes were stripped from blotted membranes by two rounds of washing in  $0.1\times\text{SSC}/0.5\%$  SDS at  $95^{\circ}\text{C}$  for 20 min and autoradiographed to ensure all probe DNA had been removed.

### 5.15. Data analysis

Sequence analysis and primer design were performed using programs from the Wisconsin Package available from the Genetics Computer Group via the Human Genome Mapping Project – Resource Centre (HGMP-RC) at the Sanger Centre, Hinxton, UK. In particular the following programs were extensively used: PileUp (sequence alignment, ‘treeplots’), BLAST (sequence similarity), REVERSE/BESTFIT/GAP (complete sequence deduction), TRANSLATE (amino acid sequence deduction) and MAP (restriction site mapping), along with the Sequence Retrieval System (SRS) for general sequence information.

## Acknowledgements

We thank Profs Vederas (University of Alberta) and Hutchinson (University of Wisconsin) for the kind provision of pTpks-100 (plasmid containing *LNKS*); Prof. Yoder (Cornell University) for the kind provision of *C. heterostrophus* and the T-toxin encoding plasmid p627PKSCH-28; and Prof. P.F. Leadlay of the University of Cambridge for *ery* clones. Dr Frank vanMiddlesworth and Dr Barrie Wilkinson of GlaxoWellcome are thanked for stimulating and helpful discussions. BBSRC and GlaxoSmithKline are kindly acknowledged by T.P.N. for the provision of a CASE award. Frank Glod is thanked for technical assistance.

## References

- [1] D.A. Hopwood, Genetic contributions to understanding polyketide synthases, *Chem. Rev.* 97 (1997) 2465–2497.
- [2] D.A. Hopwood, Forty years of genetics with *Streptomyces*: from in vivo through in vitro to in silico, *Microbiology* 145 (1999) 2183–2202.
- [3] F. Malpartida, D.A. Hopwood, Molecular-cloning of the whole biosynthetic-pathway of a *Streptomyces* antibiotic and its expression in a heterologous host, *Nature* 309 (1984) 462–464.
- [4] C. Bisang, P.F. Long, J. Cortes, J. Westcott, J. Crosby, A.-L. Matharu, R.J. Cox, T.J. Simpson, J. Staunton, P.F. Leadlay, A chain initiation factor common to both modular and aromatic polyketide synthases, *Nature* 401 (1999) 502–505.
- [5] S.E. Hallam, F. Malpartida, D.A. Hopwood, Nucleotide sequence, transcription and deduced function of a gene involved in polyketide antibiotic synthesis in *Streptomyces coelicolor*, *Gene* 74 (1988) 305–320.
- [6] F. Malpartida, S.E. Hallam, H.M. Kieser, H. Motamedi, C.R. Hutchinson, M.J. Butler, D.A. Sugden, M. Warren, C. McKillop, C.R. Bailey, G.O. Humphreys, D.A. Hopwood, Homology between *Streptomyces* genes coding for synthesis of different polyketides used to clone antibiotic biosynthetic genes, *Nature* 325 (1987) 818–821.
- [7] K.J. Stutzman-Engwall, C.R. Hutchinson, Multigene families for anthracycline antibiotic production in *Streptomyces peucetius*, *Proc. Natl. Acad. Sci. USA* 86 (1989) 3135–3139.
- [8] A. Grimm, K. Madduri, A. Ali, C.R. Hutchinson, Characterization of the *Streptomyces peucetius* ATCC-29050 genes encoding doxorubicin polyketide synthase, *Gene* 151 (1994) 1–10.
- [9] T. Schwecke, J.F. Aparicio, I. Molnar, A. König, L.E. Khaw, S.F. Haydock, M. Oliynyk, P. Caffrey, J. Cortes, J.B. Lester, G.A. Bohm, J. Staunton, P.F. Leadlay, The biosynthetic gene-cluster for the polyketide immunosuppressant rapamycin, *Proc. Natl. Acad. Sci. USA* 92 (1995) 7839–7843.
- [10] I. Molnar, T. Schupp, M. Ono, R.E. Zirkle, M. Milnamow, B. Nowak-Thompson, N. Engel, C. Toupet, A. Stratmann, D.D. Cyr, J. Grolach, J.M. Mayo, A. Hu, S. Goff, J. Schmid, J.M. Ligon, The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from *Sorangium cellulosum* So ce90, *Chem. Biol.* 7 (2000) 97–109.
- [11] N.P. Keller, T.M. Hohn, Metabolic pathway gene clusters in filamentous fungi, *Fungal Genet. Biol.* 21 (1997) 17–29.
- [12] J.B. Spencer, P.M. Jordan, Purification and properties of 6-methylsalicylic acid synthase from *Penicillium patulum*, *Biochem. J.* 288 (1992) 839–846.
- [13] J. Beck, S. Ripka, A. Siegner, E. Schiltz, E. Schweizer, The multifunctional 6-methylsalicylic acid synthase gene of *Penicillium patulum*

- its gene structure relative to that of other polyketide synthases, Eur. J. Biochem. 192 (1990) 487–498.
- [14] D.J. Bedford, E. Schweizer, D.A. Hopwood, C. Khosla, Expression of a functional fungal polyketide synthase in the bacterium *Streptomyces coelicolor* A3, J. Bacteriol. 177 (1995) 4544–4548.
- [15] J.T. Kealey, L. Liu, D.V. Santi, M.C. Betlach, P.J. Barr, Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts, Proc. Natl. Acad. Sci. USA 95 (1998) 505–509.
- [16] G.H. Feng, T.J. Leonard, Culture conditions control expression of the genes for aflatoxin and sterigmatocystin biosynthesis in *Aspergillus parasiticus* and *A. nidulans*, Appl. Environ. Microbiol. 64 (1998) 2275–2277.
- [17] I. Fujii, Y. Ono, H. Tada, K. Gomi, Y. Ebizuka, U. Sankawa, Cloning of the polyketide synthase gene *Atx* from *Aspergillus terreus* and its identification as the 6-methylsalicylic acid synthase gene by heterologous expression, Mol. Gen. Genet. 253 (1996) 1–10.
- [18] M. Walsh, Ph.D. Thesis, University of Bristol, Bristol, 1993.
- [19] Y. Takano, Y. Kubo, K. Shimizu, K. Mise, T. Okuno, I. Furusawa, Structural analysis of *Pks1*, a polyketide synthase gene involved in melanin biosynthesis in *Colletotrichum lagenarium*, Mol. Gen. Genet. 249 (1995) 162–167.
- [20] J.H. Yu, T.J. Leonard, Sterigmatocystin biosynthesis in *Aspergillus nidulans* requires a novel type-I polyketide synthase, J. Bacteriol. 177 (1995) 4792–4800.
- [21] M.E. Mayorga, W.E. Timberlake, The developmentally regulated *Aspergillus-Nidulans wA* gene encodes a polypeptide homologous to polyketide and fatty-acid synthases, Mol. Gen. Genet. 235 (1992) 205–212.
- [22] G. Yang, M.S. Rose, B.G. Turgeon, O.C. Yoder, A polyketide synthase is required for fungal virulence and production of the polyketide T-toxin, Plant Cell 8 (1996) 2139–2150.
- [23] V.A. Vinci, US Pat. 5 744 (1998) 350.
- [24] J. Kennedy, K. Auclair, S.G. Kendrew, C. Park, J.C. Vederas, C.R. Hutchinson, Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis, Science 284 (1999) 1368–1372.
- [25] L. Hendrickson, C.R. Davis, C. Roach, D.K. Nguyen, T. Aldrich, P.C. McAda, C.D. Reeves, Lovastatin biosynthesis in *Aspergillus terreus*: characterization of blocked mutants, enzyme activities and a multifunctional polyketide synthase gene, Chem. Biol. 6 (1999) 429–439.
- [26] R.H. Proctor, A.E. Desjardins, R.D. Plattner, T.M. Hohn, A polyketide synthase gene required for biosynthesis of fumonisin mycotoxins in *Gibberella fujikuroi* mating population A, Fungal Genet. Biol. 27 (1999) 100–112.
- [27] L.E.H. Bingle, T.J. Simpson, C.M. Lazarus, Ketosynthase domain probes identify two subclasses of fungal polyketide synthase genes, Fungal Genet. Biol. 26 (1999) 209–223.
- [28] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.
- [29] J.P. Jost, J.P. Saluz, DNA Methylation: Molecular Biology and Biological Significance, Birkhauser Verlag, Basel, 1993.
- [30] T.E. Cleveland, D. Bhatnagar, Individual reaction requirements of 2 enzyme activities, isolated from *Aspergillus parasiticus*, which together catalyze conversion of sterigmatocystin to aflatoxin-B1, Can. J. Microbiol., 33 (1987) 1108–1112.
- [31] J.J. Yu, P.K. Chang, G.A. Payne, J.W. Cary, D. Bhatnagar, T.E. Cleveland, Comparison of the *omtA* genes encoding *O*-methyltransferases involved in aflatoxin biosynthesis from *Aspergillus parasiticus* and *Aspergillus flavus*, Gene 163 (1995) 121–125.
- [32] R.M. Kagan, S. Clarke, Widespread occurrence of 3 sequence motifs in diverse *S*-adenosylmethionine-dependent methyltransferases suggests a common structure for these enzymes, Arch. Biochem. Biophys. 310 (1994) 417–427.
- [33] J. Vidgren, L.A. Svensson, A. Liljas, Crystal-structure of catechol *O*-methyltransferase, Nature 368 (1994) 354–358.
- [34] D.V. Santi, M.A. Siani, B. Julien, D. Kupfer, B. Roe, An approach for obtaining perfect hybridization probes for unknown polyketide synthase genes: a search for the epothilone gene cluster, Gene 247 (2000) 97–102.
- [35] T.A. Brown, Molecular Biology Labfax, 2nd edn., vol. 1, Academic Press, London, 1988.
- [36] M.J.F. Lozano, L.L. Remsing, L.M. Quiros, A.F. Brana, E. Fernandez, C. Sanchez, C. Mendez, J. Rohr, J.A. Salas, Characterization of two polyketide methyltransferases involved in the biosynthesis of the antitumor drug mithramycin by *Streptomyces argillaceus*, J. Biol. Chem., 275 (2000) 3065–3074.
- [37] G.L. Challis, J. Ravel, C.A. Townsend, Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains, Chem. Biol. 7 (2000) 211–224.
- [38] J. Hanahan, J. Mol. Biol. 166 (1983) 557–580.