Mol Gen Genet (1994) 243:506–514

ORIGINAL PAPER

Alan P. Maloney · Hans D. VanEtten

A gene from the fungal plant pathogen Nectria haematococca that encodes the phytoalexin-detoxifying enzyme pisatin demethylase defines a new cytochrome P450 family

Received: 17 May 1993 / Accepted: 30 November 1993

Abstract The gene *PDA*T9 from the fungus *Nectria haematococca* encodes pisatin demethylase, an enzyme that detoxifies the phytoalexin pisatin, an antimicrobial compound produced by pea in response to infection by this plant pathogen. *PDA*T9 was found to contain an open reading frame (ORF) encoding 515 amino acids and four introns of 52–58 nucleotides each within its coding region. The amino acid sequence F-G-A-G-S-R-S-C-I-G, indicative of the "fifth ligand binding site" present in all cytochrome P450s, occurs as residues 446 to 455, confirming that *PDA*T9 is a cytochrome P450. The deduced amino acid sequence is distinct from all other reported cytochrome P-450s, and *PDA*T9 has been assigned to a new cytochrome P450 family, CYP57. A 1.3 kb *Sac*I fragment of the *PDA*T9 ORF that lacked the fifth ligand binding site, hybridized to unique DNA fragments in *N. haematococca* isolates known to possess *PDA* genes that encode different whole cell phenotypes for pisatin demethylating activity. These genes were also tentatively identified as cytochrome P450s by the hybridization of the same fragments to separate subclones of *PDA*T9, one of which contained the fifth ligand sequence. That probe also hybridized to DNA other than that attributed to pisatin demethylase genes; these other DNAs are presumed to represent other cytochrome P450s.

Key words Phytoalexin · Detoxification · Pisatin Fungal cytochrome P450 · Plant pathogen

Communicated by C. van den Hondel

A. Maloney · H. D. VanEtten[1] (✉)
Department of Plant Pathology, Cornell University Ithaca, NY 14853, USA

[1] *Present address*:
Department of Plant Pathology, University of Arizona, Tucson, AZ 85721, USA

Offprint requests to: H.D. VanEtten

Introduction

Although cytochrome P450s are claimed to be the most versatile biological catalysts known (Porter and Coon 1991), all genes encoding cytochrome P450s are now considered members of one "superfamily" because of the high degree of conservation at the amino acid level (Nelson and Strobel 1987; Nebert et al. 1991). Pisatin demethylase is an inducible substrate-specific cytochrome P450 found in the fungus *Nectria haematococca* Berk. and Br. (anamorph *Fusarium solani*), a pathogen of pea (*Pisum sativum* L.). The demethylation of the pea phytoalexin pisatin catalyzed by this enzyme detoxifies this plant antibiotic and a major interest in this cytochrome P450 relates to its role in the pathogenicitiy of *N. haematococca* on pea (VanEtten et al. 1989).

Conventional genetic studies (Kistler and VanEtten 1984a, b; Mackintosh et al. 1989; Miao and VanEtten 1992; P.S. Mathews, unpublished data) have characterized six different gene loci (*PDA* genes) for three different phenotypes of pisatin demethylating activity (Pda). These Pda phenotypes can be distinguished by the lag period for induction and the resulting amount of enzyme activity induced: Pda^{SH} = short lag, high activity; Pda^{SM} = short lag, moderate activity: Pda^{LL} = long lag, low activity. Only fungal isolates with the Pda^{SH} or the Pda^{SM} phenotypes can cause disease on pea, implying that rapid and efficient detoxification of this plant antibiotic is required for pathogenicity. This observation and other data (Schäfer et al. 1989) have led to the hypothesis that the evolution of a specific cytochrome P450 in this fungus has allowed it to parasitize pea (VanEtten et al. 1989).

A gene encoding Pda has been cloned from a Pda^{SH} isolate (isolate T-9) of *N. haematococca* by detecting its expression in *Aspergillus nidulans* (Weltring et al. 1988). Although previous biochemical data indicate that pisatin demethylase is a cytochrome P450 (Matthews and VanEtten 1983), it has not been demonstrated that the cloned gene (called *PDA*T9) encodes this cytochrome

P450. The present study confirms this assumption by demonstrating the presence in *PDA*T9 of a conserved region characteristic of all cytochrome P450s. Furthermore, the structural characterization of *PDA*T9 indicates that this gene identifies a new family of cytochrome P450s and suggests that the previously identified *PDA* genes make up a family of closely related cytochrome P450s.

Materials and methods

Subcloning and nucleic acid purification

DNA manipulations followed conventional protocols (Sambrook et al. 1989; Birnboim and Doly 1979). The 3.3 kb *Xho*I-*Bam*HI fragment of *N. haematococca* DNA containing *PDA*T9 was isolated from pDA13 (Weltring et al. 1988) and inserted into the *Sal*I-*Bam*HI site of the pBS vector (Stratagene). The resulting plasmid, pDM1, was mapped using six-base restriction enzymes; the insert was subcloned in pBS. Fungal RNA was purified by the method of Kurtz and Lundquist (1984).

Cultures and media

Plasmids were maintained in *Escherichia coli* strain DH5α and grown on Luria-Bertani medium (Sambrook et al. 1989). All isolates of *N. haematococca* mating population (MP) VI used in this study contained active alleles at only one *PDA* locus, except for isolate T-9 and the Pda^- isolate, 44–100 (Kistler and VanEtten 1984b). The reference isolates and the genes they contained were 77-2-3 (*PDA*1), 96-17 (*PDA*2), 66-22 (*PDA*3) (Kistler and VanEtten 1984b), 196-10-7 (*PDA*4) (Mackintosh et al. 1989), 159-5-3 (*PDA*5) (P. Matthews, unpublished), 156-30-6 (*PDA*6-1) and T200 (*PDA*6-2). *PDA*6-1 and *PDA*6-2 are two different alleles at the *PDA*6 locus (Miao and VanEtten 1992). Genes *PDA*1 and *PDA*5 encode a Pda^{SH} phenotype and *PDA*4 a Pda^{SM} phenotype. All the other *PDA* genes result in a Pda^{LL} phenotype. Isolate T9 has a Pda^{SH} phenotype but the number of active *PDA* alleles in this isolate has not been determined by conventional genetics because of its poor fertility in crosses (Kistler and VanEtten 1984b). Isolates were grown and maintained as described previously (Kistler and VanEtten 1984b).

Pisatin demethylase and *PDA*T9 mRNA induction

Pisatin demethylase activity and *PDA*T9 mRNA biosynthesis were induced by the addition of pisatin to washed fresh mycelium in phosphate buffer (Kistler and VanEtten 1984b). At 1 h intervals, mycelium from 15 ml aliquots was frozen for subsequent RNA purification and enzyme activity was measured by adding radiolabeled pisatin to an additional 5 ml aliquot. Pisatin demethylase activity was determined by Method 1 of Kistler and VanEtten (1984b) Enzyme activity peaked at 6 h; mRNA isolated from pooled 5, 6, and 7 h samples was used for all RNA experiments.

Northern analysis

N. haematococca RNA was electrophoretically fractionated on formaldehyde gels, transferred to Genescreen Plus (Dupont), and hybridized using standard protocols (Sambrook et al. 1989). Ribosomal RNA and alkaline-denatured DNA restriction fragments of known sizes were used as size standards.

RNase protection experiments

Transcript sizes and endpoints were determined by RNase protection assays (Winter et al. 1985). Single-stranded radiolabeled RNA probes were produced using T7 or T3 RNA polymerase from subclones of the pDM1 3.3 kb insert, and were uniformly labeled with [α^{32}P]UTP (Amersham; >400 Ci/mmol). The templates were truncated at a restriction site within the insert or at the vector insert junction distal to the appropriate promoter to produce homogeneous 3′ ends and to reduce contamination by RNA initiated at an inappropriate promoter. Reaction mixtures of approximately 5×10^5 cpm of probe, 5 mg of fungal RNA and 25 mg of yeast tRNA (carrier) were denatured at 85° C, annealed overnight at 50° C, and digested at 22° C for 12, 60, and 120 min using 12 μg/ml RNase T_1 and 0.67 μ g/ml RNase A. Reaction products were electrophoretically fractionated on a 6% acrylamide, 7 M urea gel and autoradiographed.

Primer extension and RNA sequencing

An anti-sense oligonucleotide primer (5′-AAGTGGATGAGTGTAAGA-3′) starting at nucleotide (nt) −23 was used to initiate the direct sequencing of the mRNA 5′ end. Primer extension and dideoxy mRNA sequencing were carried out as described by Traynor and Levings (1986). The products were fractionated on a 6% polyacrylamide, 7 M urea gel and autoradiographed.

cDNA cloning

Fungal mRNA was obtained from pisatin-induced fungal mycelium by oligo(dT) cellulose column chromatography of bulk cellular RNA (Jacobson 1987). A cDNA library was constructed using oligo(dT) priming, reverse transcriptase, and Klenow fragment. The products were ligated into the *Hin*cII site of pBS, transformed into DH5α and screened by colony hybridization with the pDM1 insert. Although 1.6–1.8 kb cDNAs homologous to the pDM1 insert (putative full-length cDNAs) were detected by Southern hybridization before ligation, only half-length 3′ cDNA clones were obtained in plasmid vectors. One of the plasmid clones, named A1, was sequenced.

To obtain a cDNA from the upstream half of the mRNA, two oligonucleotide primers were synthesized to direct the amplification of the upstream portion of the *PDA*T9 cDNA via the polymerase chain reaction (PCR). One 25 nt primer (5′-CCGGATCCACCACTCATCTCTCTCT-3′), included 20 nt of the 5′ end of the mRNA as determined by the primer extension studies described above, and a *Bam*HI restriction site adaptor. The second primer, a 20mer (5′-CTGGATGATTCGTCGCTCGA-3′), was complementary to a site 56 nt downstream of the 5′ end of cDNA clone A1. These primers were used to amplify *N. haematococca* T-9 cDNA (Saiki et al. 1988) producing approximately 0.75 μg of a 1 kb fragment of DNA, which was digested with *Bam*HI and ligated into the *Sma*I and *Bam*HI sites of pBS. Multiple clones with identical restriction patterns were obtained (data not shown). One of these, cDNA 67-1, was sequenced.

Sequencing and sequence analysis

Genomic DNA and cDNA subclones were sequenced using modified T7 DNA polymerase (Sequenase, USB) according to the instructions of the manufacturer. By using overlapping segments, the entire original 3.3 kb fragment was sequenced multiple times and in both directions. The sequences of the genomic and cDNA clones were assembled using the Microgenie program (Queen and Korn 1984).

The amino acid sequence of the *PDA*T9 pisatin demethylase was deduced from a comparison of the cDNA and genomic sequences. The coding region of the *PDA*T9 cDNA was analyzed using the GCG sequence analysis package (Devereux et al. 1984). Pairwise comparisons were carried out between *PDA*T9 and representative sequences from established families of mammalian, fungal, and bacterial cytochrome P450s. Comparisons using the Gap and Best Fit programs of the GCG sequence analysis pack-

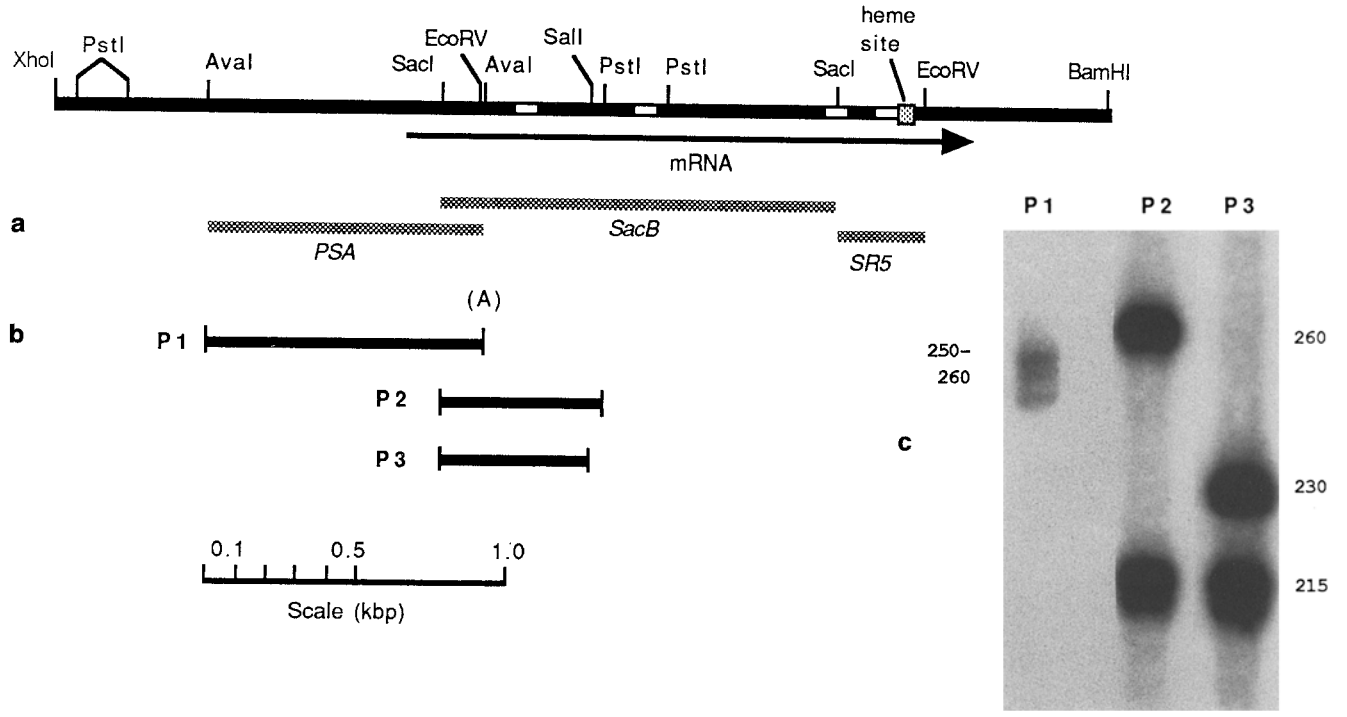

Fig. 1a–c Restriction site and transcript maps of the approximately 3.3 kb *Xho*I-*Bam*HI *Nectria haematococca* DNA insert of clone pDM1. Restriction sites delineate some of the DNA (**a**) and RNA (**b**) probes used to map the *PDA*T9 transcript. Transcription direction and approximate position of the transcript are shown by the *arrow*. **a** The three DNA probes used in Southern analysis in this study. **b** Three of the in vitro-synthesized RNA probes used in RNase protection experiments (as described in text), and **c** the resulting electrophoretic patterns of RNase-treated hybrids of those single stranded probes and RNA from pisatin-induced mycelium of *N. haematococca* isolate T9. The 3′ end of probe P1 is an *Ava*I site (A). Numbers to the sides of the autoradiogram panels indicate approximate sizes, in nucleotides, of the reaction products. Samples in lanes P2 and P3 were electrophoresed on the same gel

Gap and Best Fit programs of the GCG sequence analysis package resulted in equivalent pairwise sequence alignments and sequence-similarity values. The algorithm incorporated into the Gap program (Needleman and Wunsch 1970; Gotoh 1982) has been used to study the relationships among these enzymes and to classify the cytochrome P450s into families (Nelson and Strobel 1987).

Southern analysis

Fungal genomic DNA was purified following the procedure described by Turgeon et al. (1985); DNA was precipitated with 1.0 M NaCl 8.2% polyethylene glycol. The DNA was digested with restriction endonucleases, electrophoretically fractionated according to standard procedures (Sambrook et al. 1989), and transferred to Genescreen Plus (Dupont) or Hybond N (Amersham) for hybridization analysis. Southern hybridizations and washes were carried out as recommended for Genescreen Plus by the manufacturer.

Results

RNA mapping

Although isolate T9 contains both a Pda^{SH} and a Pda^{LL} gene (see below), only the Pda^{SH} gene (*PDA*T9) is expressed under the conditions used in this study (K. Hirschi, unpublished). In Northern hybridizations, a single 1.6 kb RNA was detected when hybridized with each of three subclones containing contiguous *Sac*I fragments of pDM1 (data not shown), and the middle fragment (SacB) gave the stronger signal, indicating that the mRNA is transcribed from near the middle of the pDM1 insert (Fig. 1a and map).

RNA transcribed in the *Bam*HI-*Xho*I direction was protected from RNase digestion in the RNase protection assays. In contrast, RNA transcribed in the *Xho*I-*Bam*HI direction was not (data not shown) indicating the 5′ end of the mRNA lay toward the *Xho*I site (as illustrated in Fig. 1). With probe P1 (Fig. 1b, 1c), three protected fragments were detected, which differed in length by about 5 nt. The 5′ ends of these fragments lay approximately 125 to 135 nt upstream of the first *Sac*I site.

The precise location of the mRNA 5′ end was determined from primer extension assays and direct RNA sequencing of the 5′ region, using an oligonucleotide primer. Six distinct primer extension products were seen within the sequence 5′-CACCACUCAUCUCUCU-CUCAUUUCUCGCUA-3′ (Fig. 2), in which the six inferred 5′ ends of the mRNA are underlined. The first of the underlined cytosine residues in this sequence lies 138 nt upstream from the 5′-most *Sac*I site of the pDM1 insert. The multiple primer extension products correlat-

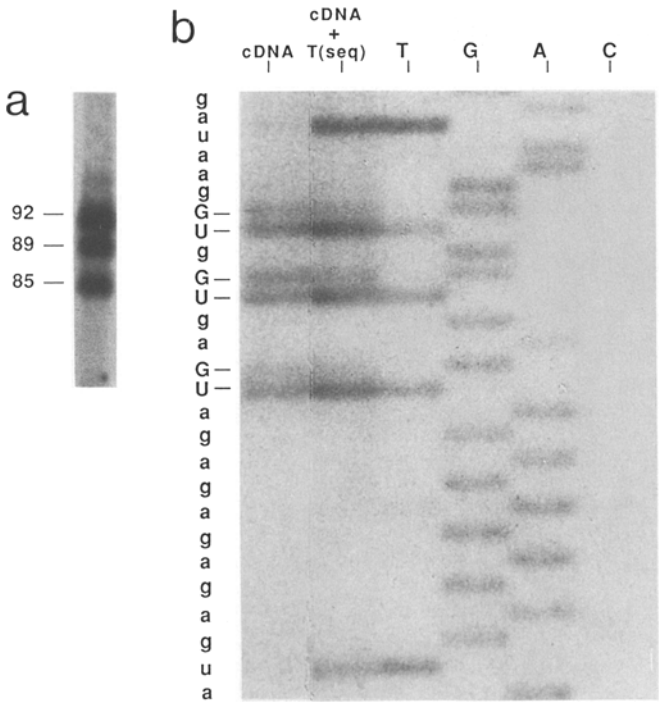

Fig. 2a, b Primer extension assay of mRNA 5′ end. **a** Autoradiogram of the cDNA products of a primer extension experiment electrophoresed on a denaturing polyacrylamide gel. **b** Autoradiogram of the full-length primer extension products and the DNA sequence generated with the same primer, electrophoresed alongside each other. Labels above the last four lanes correspond to the dideoxynucleotide used to generate the reaction products in that lane. The second lane from left shows the primer extension products co-electrophoresed with sequencing reaction products of the dideoxy-T reaction

ed with the multiple protected fragments in the RNase protection assays that used probe P1 (Fig. 1c) and are consistent with multiple transcription start sites as seen in the pectin lyase gene from this fungus (Gonzalez-Candelas and Kolattukudy 1992) and other filamentous fungi (Ballance 1991).

RNase protection assays were performed with probes P2 and P3 (Fig. 1b) to determine whether introns existed between the first *Sac*I site and the *Pst*I site 535 nt downstream. Probe P2, transcribed from the *Sac*I-*Pst*I (535 nt) fragment, resulted in 215 nt and 260 nt fragments and probe P3, transcribed from the *Sac*I-*Sal*I fragment (502 nt), resulted in 215 nt and 230 nt fragments (Fig. 1c). The sum of the lengths of protected fragments for each probe was approximately 50–60 nt shorter than the probes themselves, indicating the presence of a ca. 50 nt intron located 215 nt downstream of the *Sac*I site. Similar analysis, using other probes that were transcribed in vitro from other subclones of *PDA*T9, identified three other regions of mismatch between the length of a probe and the sum of the fragments it protected. One of these putative introns was near the middle of the *Sac*B fragment of pDM1, and the other two were near the second *Sac*I site in the downstream portion of the insert (data not shown).

Genomic and cDNA clone sequence

The pDM1 *Xho*I-*Bam*HI insert was sequenced and found to be 3276 nt long. While the normal consensus transcription-control sequences of eukaryotic genes, the TATAA and CAAT boxes found near the transcription initiation sites, do not always occur in genes of filamentous fungi (Ballance 1991), a TATAAATA sequence occurs approximately 40 nt upstream from the 5′ end of *PDA*T9 mRNA (Fig. 3). A similar sequence has been found in a similar location in the cutinase and pectin lyase genes of *N. haematococca* (Soliday et al. 1989; Gonzalez-Candelas and Kolattukudy 1992).

The sequence of cDNA clones A1 and 67-1 when compared with the genomic sequence allowed identification of four introns of 52 to 58 nt (Fig. 3), a common size for introns of filamentous fungi (Ballance 1991). The intron/exon boundary sequences are identical to, or differ by a single base from, the consensus sequences for intron/exon boundaries in other filamentous fungi (Ballance 1991). The location of a polyadenylation site was determined directly from the sequence of A1 (Fig. 3).

A single open reading frame of 515 amino acids, beginning with a methionine codon, was located in the cDNA sequence. The sequence F-G-A-G-S-R-S-C-I-G was located at positions 446 to 455 (Fig. 3). This sequence is highly conserved in cytochrome P450s with the underlined amino acid residues being strictly conserved in nearly every cytochrome P450 so far examined (Gonzalez 1988). The cysteine in this sequence constitutes the proximal thiolate ligand to the iron atom of heme, also called the "fifth ligand binding site" (Gotoh and Fujii-Kuriyama 1989). The presence of this diagnostic sequence confirms that *PDA*T9 encodes a cytochrome P450.

Comparison of *PDA*T9 amino acid sequence with other cytochrome P450s

Comparisons were made between the open reading frame of *PDA*T9 and the amino acid sequences of 32 cytochrome P450s representing 11 of the gene families (14 subfamilies) of the cytochrome P450 superfamily (Nebert et al. 1991). In all of the comparisons, the fifth ligand binding sites of the two sequences were aligned. Sequence identity and sequence similarity scores calculated by the GCG program (using a gap weight of 3.0 and a gap length weight of 0.1) indicated that the maximum sequence similarity between *PDA*T9 and other cytochrome P450s was 52.2%, but most of the values were below 50%. The maximum amino acid sequence identity was 26.5%. An amino acid sequence identity level of 40% has been defined as the minimum required to include two P450s in the same family (Nelson and Strobel 1987; Nebert et al. 1991). Therefore, *PDA*T9 is distinct from all other cytochrome P450s so far described, including those of other filamentous fungi (Nelson et al. 1993) and has been assigned, as the first member, to a

```
-300 AGCTTTTCTGTCCAAAACTCTCCTGAATCCCATCTCCTCAAGTTGGTCATTCACACGGGTTGGTTGGCTTAGCCTTCCCTTTTGCGATTACTACGCCTCA
-200 TGGCAGATACAAGCCCTGAAAAATTCTTAGTGCGTGAGGGGTTGCCTAGGTAGAGTAATCTTTAGTTATAAATACCCATCCCTCCTTGTTATCGTGACTA
-100 TTCCACCACTCATCTCTCTCTCATTTCTCGCTATCGACAGTGACTTGTCGGCTCGAATCATATCTCGATCTCCTGGCTCTTACACTCATCCACTTCAAAC
   1 ATGCTGGTAGACACTGGTCTGGGGCTCATCAGCGAGCTCCAAGCCAAACTTGGCTGGGCTGTCCTCCTTCAGATCGTCCCTATCACCATTGTCGCCTACA
   1  M  L  V  D  T  G  L  G  L  I  S  E  L  Q  A  K  L  G  W  A  V  L  L  Q  I  V  P  I  T  I  V  A  Y
 101 ATCTTCTCTGGTTCATCTACGCGTCCTTCTTTTCGAGTCTGAGAAAGATCCCTGGTCCTTTTCTTGCGCGGATATCCCGAGTCTGGGAGATGAAAAAGAC
  34 N  L  L  W  F  I  Y  A  S  F  F  S  S  L  R  K  I  P  G  P  F  L  A  R  I  S  R  V  W  E  M  K  K  T
 201 GGCCACTGGAAACATCCACGAGATCATGATGGATCTACACCGACGTCACGGTGAGTCTGTGGCTTGGAATTGGCAGTCCTTATACTTACCATATCTTTTT
  68    A  T  G  N  I  H  E  I  M  M  D  L  H  R  R  H
 301 ACCAGGAGCCATCGTTCGAATAGGACCGAGGCGGTACGACTTTGACACCATGGAAGCTCTCAAGATCATCTACCGCATTGGCAACGCCCTCCCCAAAGCC
  84       G  A  I  V  R  I  G  P  R  R  Y  D  F  D  T  M  E  A  L  K  I  I  Y  R  I  G  N  A  L  P  K  A
 401 GACTACTATAAACCCTTTGGCCTGCCGTCGTTCCCGAATCTTTTCGACGAGCAAAACCCCGCTCGTCATTCAGCGATAAAGAAGCAGGTTGCCTCTCTGT
 116  D  Y  Y  K  P  F  G  L  P  S  F  P  N  L  F  D  E  Q  N  P  A  R  H  S  A  I  K  K  Q  V  A  S  L
 501 ACACCATGACGGCGTTGCTATCCTACGAAGAGGGGGTCGACGGCCAGACGGCCATCCTGAAGGAGCAGCTGCAGAGGTTCTGCGACCAGAAGCAGGTCAT
 149 Y  T  M  T  A  L  L  S  Y  E  E  G  V  D  G  Q  T  A  I  L  K  E  Q  L  Q  R  F  C  D  Q  K  Q  V  I
 601 TGACCTGCCTCGATTCCTCCAGTACTATGCTTTTGACGTCATTGGAGTTATTACGGTGAGTTCACTGTCACAAGCGAGCCGTGGAGCATTATCTGACATG
 183    D  L  P  R  F  L  Q  Y  Y  A  F  D  V  I  G  V  I  T
 701 TGTTTCTGCTAGGTTGGCAAGTCGATGGGCATGATGGAGTCGAACTCGGATACAAACGGTGCCTGCAGTGCCCTGGACGGCATGTGGCACTACGCTTCCA
 201              V  G  K  S  M  G  M  M  E  S  N  S  D  T  N  G  A  C  S  A  L  D  G  M  W  H  Y  A  S
 801 TGATGGCTTATATACCAAACATGCACGCCTGGTGGCTGCGTCTCTCTTCACTCCTACCGATCGAAGTCCCCATCAAGGGCCTGACCGAGTATGTCGAGCG
 230 M  M  A  Y  I  P  N  M  H  A  W  W  L  R  L  S  S  L  L  P  I  E  V  P  I  K  G  L  T  E  Y  V  E  R
 901 ACGAATCATCCAGTACAGGCTCAAGGCGGCCGAGTTCGGTGACGATGCCGCACTCAAGGGCGAGAACAACTTCCTGGCCAAGCTGCTCCTCATGGAGAAG
 264   R  I  I  Q  Y  R  L  K  A  A  E  F  G  D  D  A  A  L  K  G  E  N  N  F  L  A  K  L  L  L  M  E  K
1001 AAGGGAACGGTCACTCCGGTAGAGACTCAGCAGGCTGTTGGTCTGAACATTGGGGCTGGGTCTGATACGACAGCGAATGCTCTGAGCACTATCCTGTACT
 297  K  G  T  V  T  P  V  E  T  Q  Q  A  V  G  L  N  I  G  A  G  S  D  T  T  A  N  A  L  S  T  I  L  Y
1101 ACCTGTACACAAACCCACGTACTCTGCACACTCTTCGAGAGGAGTTGGAAAGATATGTTAAGGACGGCCCCATTAGTTTTCAGCAATCGCAGAGCATGCC
 330 Y  L  Y  T  N  P  R  T  L  H  T  L  R  E  E  L  E  R  Y  V  K  D  G  P  I  S  F  Q  Q  S  Q  S  M  P
1201 TTACCTACAGGCGGTTATCAAGGAAGCCCTGCGGCTGCATCCGGGTGTGGGCACGCAACTGACGCGAGTCGTGCCAAAGGGTGGCCTGGTCATCGAGGGA
 364   Y  L  Q  A  V  I  K  E  A  L  R  L  H  P  G  V  G  T  Q  L  T  R  V  V  P  K  G  G  L  V  I  E  G
1301 CAGTTCTTCCCCGAGGGTGTAAGTCGAATCTGTCTGCCTCAAGTGTTGAGCTCAAACTAATGATGAGCAGACTGAGGTTGGTGTCAACGGATGGGCTCTA
 397  Q  F  F  P  E  G                                                          T  E  V  G  V  N  G  W  A  L
1401 TATCACAACAAAGCGATCTTTGGCAACGATGCAAGCATATTCCGCCCCGAGCGGTGGTTGGAGGCAAACGAGAATATCAACATTGGGGGATCTTTCGCGG
 413  Y  H  N  K  A  I  F  G  N  D  A  S  I  F  R  P  E  R  W  L  E  A  N  E  N  I  N  I  G  G  S  F  A
1501 TAAGTGACCTTGGCATCACAGTATTAATCGGTGGTAACTGACGCCTCTGCTGTACAGTTTGGAGCCGGTTCCCGATCCTGTATTGGCAAGAACATCAGCA
 446                                                                F  G  A  G  S  R  S  C  I  G  K  N  I  S
1601 TCTTGGAAATGTCAAAGGCCATTCCCCAGATTGTCCGGAACTTTGATATCGAGATTAACCACGGGGATATGACATGGAAGAACGAGTGTTGGTGGTTTGT
 460 I  L  E  M  S  K  A  I  P  Q  I  V  R  N  F  D  I  E  I  N  H  G  D  M  T  W  K  N  E  C  W  W  F  V
1701 CAAGCCTGAATACAAGGCCATGATTAAGCCTCGAAGATGCTGCCTATCAAGGGATGAATCCTTAGTCTAGTGTATATCTTTTGCATTCTACCACCTGTCA
 494   K  P  E  Y  K  A  M  I  K  P  R  R  C  C  L  S  R  D  E  S  L  V END                          (poly-A)
1801 AATGCACCGGCCTGGAATGTTGAAACTTTTGGACTCCTCCCTGCTACAATCAGAGGTATCCTGCACCATTGCCAGCCTCTCACAACTCCATTCTCTTGCA
```

Fig. 3 Nucleotide and amino acid sequences of *PDA*T9. The first line displays the nucleotide sequence of the genomic DNA; the second line displays the *PDA*T9 open reading frame amino acid sequence, deduced from the cDNA sequence. The nucleotide sequence numbering begins at the putative translational start codon. The TATAAATA sequence upstream of the mRNA 5′ end and the four introns are *underlined*, the 5′ endpoints of the mRNA are *double-underlined*, and the sequence of the fifth ligand binding site of cytochrome P450 is in *bold*. The nucleotide sequence presented here is from nucleotides −300 to +1900 of the original 3.3 kb pDM1 clone. The complete sequence, extending from the *Xho*I site (−1189) to the *Bam*HI site (2087) of the pDM2 insert, is filed with Genbank (Accession no. L20976)

new family of cytochrome P450s, CYP57 (Nebert et al. 1991).

Southern analysis of *N. haematococca* isolates

Three regions of *PDA*T9 were used as probes in Southern hybridizations to characterize the *N. haematococca* isolates and the relationship between *PDA*T9 and the other *PDA* genes. These regions were (1) the central

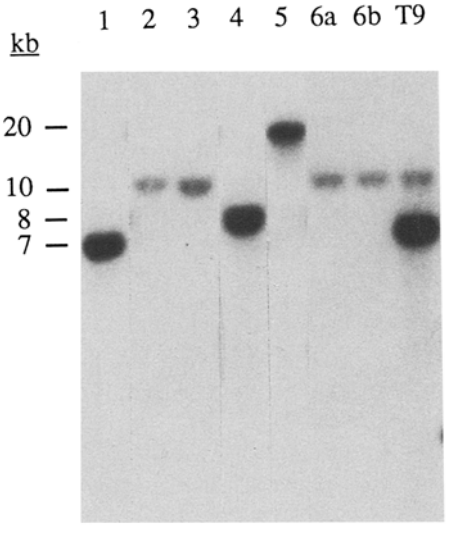

Fig. 4 Southern hybridizations of SacB to genomic DNAs from *Nectria haematococca* reference isolates digested with *Xho*I. Except for T9, the number above the lane refers to the *PDA* gene carried by the isolate whose DNA was electrophoresed in that lane. Lane 6a, PDA6-1; lane 6b, PDA6-2

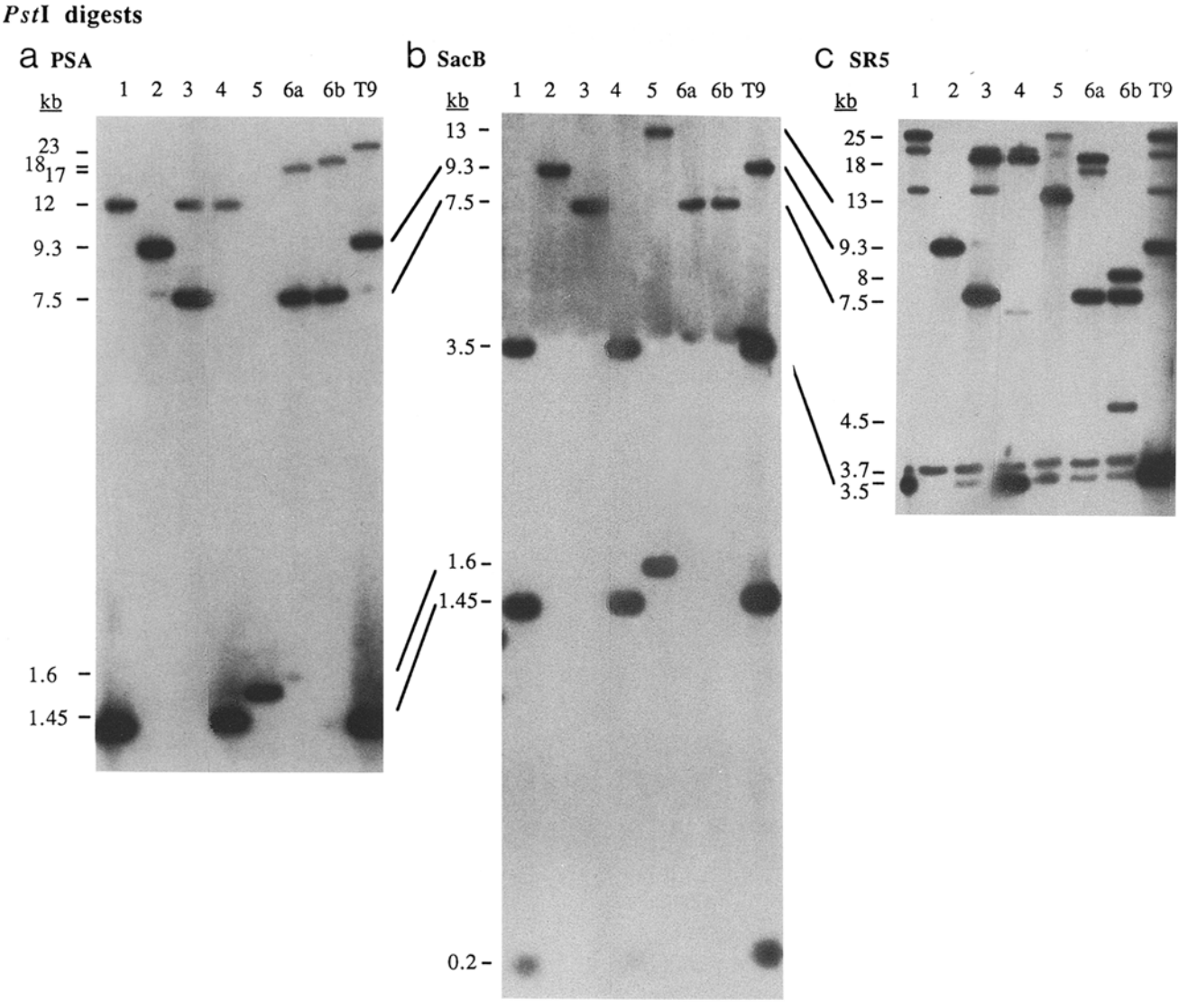

Fig. 5a–c Southern hybridization of probes PSA, SacB, and SR5 to *Pst*I-digested genomic DNA of *Nectria haematococca* reference isolates. Lane designations are as in Fig. 4. Hybridizations were to different Southern transfers, so similar-sized fragments do not appear at the same position on each panel. *Line segments* between panels indicate identical fragments detected with different probes. No fragments were detected of a size below 1.45 kb in **a** or below 3.5 kb in **c**

1.35 kb *Sac*I fragment (probe SacB) consisting of about 1.25 kb of the coding sequence of *PDA*T9, the first two introns, and a portion of the third intron; (2) a 0.9 kb *Ava*I fragment (probe PSA) consisting of the upstream end of SacB and an additional 720 nt of upstream noncoding sequence, and (3) a 318 nt *Sac*I-*Eco*RV fragment (probe SR5) containing part of intron 4 and the entire fifth ligand binding site (Fig. 1). The restriction enzymes used for genomic DNA digests were *Bam*HI, *Xho*I, *Eco*RV, *Sac*I, *Pst*I, *Pvu*I, and *Sal*I.

The results of Southern hybridizations to genomic digests were consistent with the SacB probe being a specific probe for *PDA* genes, while the other probes, in particular SR5, detected additional fragments. The autoradiograms of the Southern hybridizations shown here (Figs. 4, 5, 6) illustrate these observations as well as demonstrate how the features of *PDA*T9 can be used to characterize and distinguish different *PDA* genes. Occasionally there were fragments that showed weak hybridization to these different probes (for instance SacB hybridization to a ca. 9 kb fragment of the *Eco*RV-digested genomic DNA in lane 1 of Fig. 6b) but in the remainder of this section and in the Discussion, we will refer only to strongly hybridizing bands unless stated otherwise.

SacB failed to hybridize to *Xho*I-digested genomic DNA of the Pda$^-$ isolate 44–100 (not shown) and hybridized to a single fragment in *Xho*I-digested of the reference isolates carrying a single active gene, and to two fragments in isolate T-9 (Fig. 4). One of the fragments in isolate T-9 was the same size as that detected in isolate 77-2-3, which contains the Pda ^{SH} gene *PDA*1, and the other fragment was the same size as that detect-

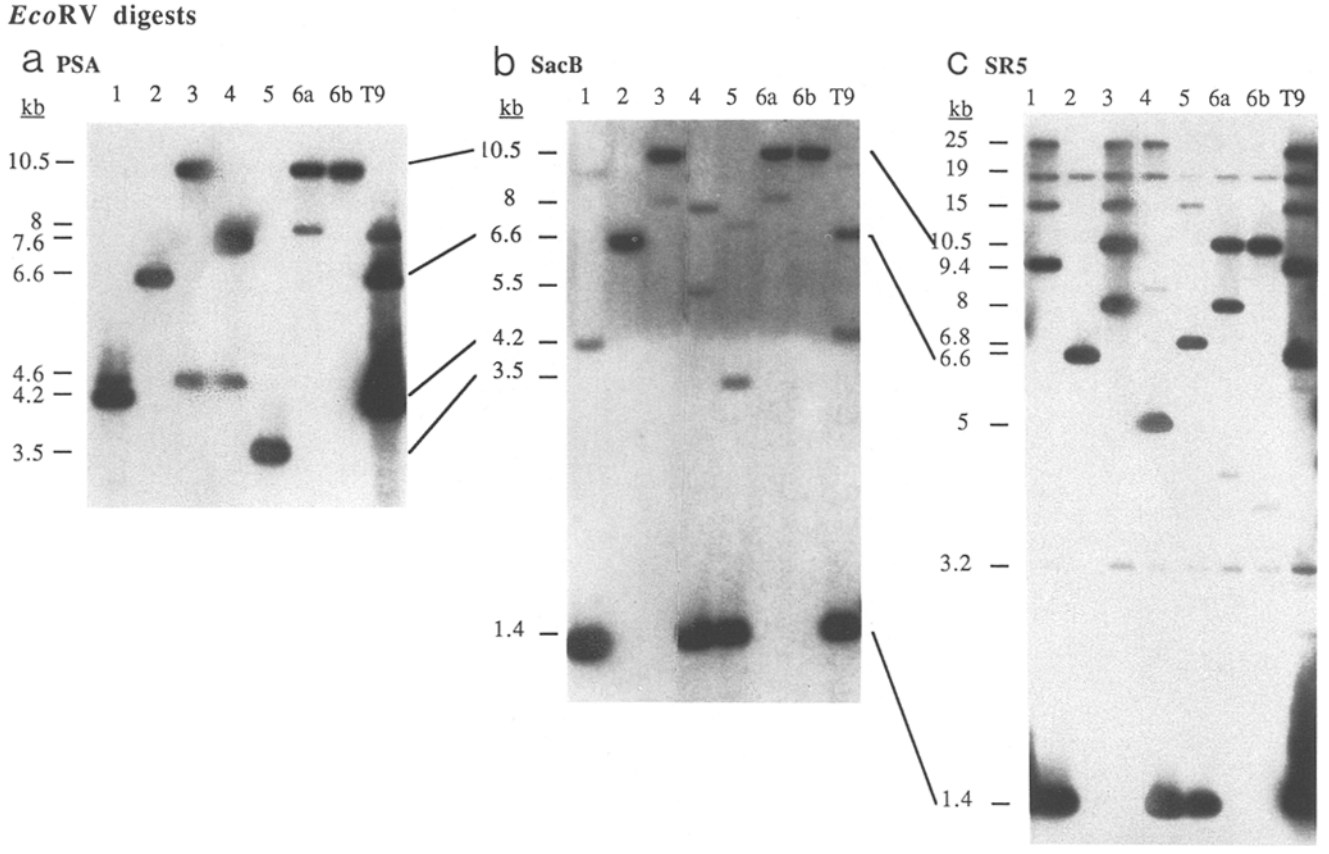

Fig. 6a–c Southern hybridization of probes PSA, SacB, and SR5 to *Eco*RV-digested genomic DNA of *Nectria haematococca* reference isolates. Lane designations and labeling of panels are as in Fig. 5. No fragments were detected of a size above 10.5 kb in **a** or **b**, or below 3.5 kb in **a**

ed in all the isolates containing Pda^{LL} genes (*PDA*2, 3, 6-1, 6-2). All subsequent analyses were consistent with isolate T-9 containing two *PDA* genes: *PDA*T9, which appeared to be a homolog of, or identical to, *PDA*1, and a *PDA* gene that was a homolog of, or identical to, *PDA*2 (e.g. Figs. 5b, 6b).

Fragments that hybridized strongly to SacB in 196-10-7, which contained *PDA*4, a Pda^{SM} gene, were similar to *PDA*1 fragments except in *Xho*I (Fig. 4), *Eco*RV (Fig. 6b) and *Pvu*I digestions (not shown). Most restriction fragments hybridizing to SacB in DNA of isolate 159-5-3, containing *PDA*5, a Pda^{SH} gene, were distinct from those of the reference isolate (77-2-3) containing the other Pda^{SH} gene (*PDA*1), and all other isolates (e.g. Figs. 4, 5b). In contrast, most of the fragments detected in isolates 62-22, 156-30-6, and T200, containing the Pda^{LL} genes *PDA*3, *PDA*6-1, and *PDA*6-2, respectively, were the same size, but distinct from those of the other Pda^{+} isolates, including isolate 96-17, which carries the fourth Pda^{LL} gene, *PDA*2 (Figs. 5b, 6b). The only restriction fragment detected by SacB that distinguished *PDA*6-1 from *PDA*3 and *PDA*6-2 was a 1.5 kb *Sac*I fragment that was absent in the *PDA*6-1 isolate but present in the other two (not shown).

Southern hybridizations performed on the same set of genomic DNAs with probes PSA and SR5 detected several DNA fragments other than those contained in the *PDA* genes themselves (Figs. 5a–c, 6a–c). For instance, from the genomic sequences of *PDA*T9 and the results of Southern hybridizations with SacB, SR5 was predicted to hybridize to a single 1.5 kb *Eco*RV fragment in 77-2-3 (*PDA*1). A fragment of this size was detected with SR5, but four additional fragments were also detected (Fig. 6c, lane 1). Similar results were obtained on hybridization of SR5 to the genomic DNA of each of the other isolates (Fig. 6c, lanes 2–6b). Hybridization with PSA resulted in bands additional to those predicted from the *PDA*T9 sequence, but these comprised a less diverse set than those hybridizing to SR5 (Figs. 5a, 6a).

Discussion

A previous study by Miao et al. (1991a) indicated that when the 3.3 kb *Xho*I-*Bam*HI fragment containing *PDA*T9 was used as a probe it detected a single *Xho*I-*Bam*HI fragment, ca. 3.3 to 23 kb, in reference isolates carrying single active *PDA* alleles and normally failed to hybridize to Pda^{-} isolates. It was proposed that the *PDA* genes are unique cytochrome P450s; this suggestion is consistent with the current sequence analysis of *PDA*T9 and the Southern hybridization analyses. In the present study, three subclones of *PDA*T9 were used as

probes on restriction digests of genomic DNA of *N. haematococca* isolates carrying different *PDA* genes; only the SacB probe, which contains a large portion of the opening reading frame of *PDA*T9, while lacking the highly conserved heme-binding fifth ligand site found in all P450s, had a specificity similar to the 3.3 kb *Xho*I-*Bam*HI fragment containing *PDA*T9 (Miao et al. 1991a vs Figs. 4–6 and data not shown). In each isolate that carried a *PDA* gene, the restriction pattern detected by SacB was simple, comprising one or two bands, and the sum of the sizes of the detected restriction fragments was consistent with the presence of a single *PDA* gene in each such isolate. We suggest therefore that the *PDA* genes in *N. haematococca* comprise a closely related family of genes, and that each of the fragments that hybridizes strongly to SacB shown in Figs. 4, 5b, and 6b contains part of the coding region of the single active *PDA* gene present in each reference isolate. In some instances, weakly hybridizing fragments were detected with SacB (e.g. Fig. 6b, lane 1); the same size fragment hybridized strongly to SR5 (e.g. Fig. 6c, lane 1) and may represent another P450 with weak sequence similarity to the *PDA* genes (see below).

The Southern hybridization analysis using the SacB probe also permitted the known *PDA* genes, with the exceptions of *PDA*3 and *PDA*6-2, to be distinguished from each other, and allowed a rough estimation of the degree of similarity between the genes; *PDA*1, *PDA*4, and, to a lesser extent, *PDA*5 appear most similar to *PDA*T9. As a group, these genes are distinct from the other *PDA* genes, based on the number of SacB-hybridizing fragments predicted from the sequence of *PDA*T9 that were held in common: *PDA*1-, *PDA*4- and *PDA*5-containing isolates all had a 1.5 kb *Eco*RV fragment (Fig. 6b); *PDA*1 and *PDA*4 isolates had the predicted 0.2 kb *Pst*I (Fig. 5b) and a 1.35 *Sac*I fragment (data not shown). Further similarities, especially between *PDA*1 and *PDA*4, were observed: fragments detected with SacB that were only partially contained in the probe sequence in *PDA*T9 were detected in *PDA*1 and *PDA*4 as well (e.g. the 1.45 kb *Pst*I fragment in *PDA*1 and *PDA*4 isolates; Fig. 5b).

In contrast, DNA of isolates containing Pda^{LL} genes (*PDA*2, *PDA*3, *PDA*6-1 and *PDA*6-2) all lacked these fragments and thus lacked several of the restriction sites that lie within the structural gene of *PDA*T9. It is of interest that these differences correlate with the different levels of whole cell activity conferred by the genes, i.e., genes conferring the Pda^{SH} and Pda^{SM} phenotype versus genes conferring the Pda^{LL} phenotypes. This restriction site comparison is too limited to allow one to infer whether the differences between the phenotypes conferred by these genes are due to regulatory differences or biochemical properties of the structural gene products, but it does suggest the presence of two structurally different groups of PDA genes in *N. haematococca*.

Although the SacB probe appeared specific for *PDA* genes, the other two subclones of *PDA*T9, probes PSA and SR5, detected more restriction fragments than could be accounted for by hybridization to *PDA* genes only. For instance, from the results with SacB (Fig. 5b) and the sequence of *PDA*T9, PSA was predicted to hybridize only to a 1.45 kb *Pst*I fragment in isolates with *PDA*1 (77-2-3), and *PDA*4 (196-10-7). Strong hybridization of PSA was also detected to a 12 kb fragment in 77-2-3 and 196-10-7 (Fig. 5a, lanes 1 and 4). Similar results were seen in comparisons of other isolates (Fig. 5a vs 5b) and in the PSA hybridization to *Eco*RV digests (Fig. 6a vs 6b). Since PSA contained a large part of the upstream region of *PDA*T9, as well as several hundred nucleotides of the 5′ terminal coding region, it is possible that the additional fragments detected with PSA represent a regulatory segment of another gene that is shared by pisatin demethylase or another gene for a membrane-bound protein whose membrane-spanning sequence is similar to that of *PDA*T9, as transmembrane peptides are located at the amino-termini of cytochrome P450s (Gotoh and Fujii-Kriyama 1989). Alternatively, the probe could simply contain a fragment of some unrelated gene.

The additional fragments detected with SR5 probably represent other P450 genes, as this probe contains the highly conserved heme binding site. It is not known how many P450s exist in filamentous fungi but all organisms are predicted to contain a number of these enzymes (Nelson and Strobel 1987). Some of the fungal P450s are involved in species-specific traits such as the catabolism of specific carbon sources, while others are involved in highly conserved and essential biochemical processes such as the synthesis of membrane sterols (Kaergel et al. 1990). The common size of the additional fragments detected by SR5 (Figs. 5c, 6c) suggests the presence of a highly conserved non-*PDA* cytochrome P450.

Previous genetic analyses have indicated that *PDA* genes occur at different loci (Kistler and VanEtten 1984b; Mackintosh et al. 1989; Miao and VanEtten 1992; P. Matthews, unpublished). However, these conclusions may not be valid. In the work described here, the Pda^{LL} genes were detected by SacB as single DNA fragments larger than 7 kb (Figs. 4, 5b, 6b). Unless the Pda^{LL} genes are unusually large, or are interrupted by many or very long introns, one or both ends of each of those large restriction fragments must reside well outside the coding regions of the genes. The lack of diversity in the size of the SacB-homologous fragments of *PDA*3, *PDA*6-1, *PDA*6-2, and to some degree *PDA*2, suggests that their immediate chromosomal environments in the different reference strains may be very similar to each other.

Furthermore, recent results indicate that *PDA*6-1 and possibly all the *PDA* genes are located on meiotically unstable chromosomes (Miao et al. 1991a, b). The basis for the original conclusion that *PDA* genes are at different loci was the recovery of Pda^- progeny from crosses between Pda^+ parents. If the chromosomes bearing these genes can be lost during meiosis, Pda^- progeny could result from a cross, even though both

parents contained a *PDA* gene at the same locus. In this case, the loss of parental phenotype in the progeny would falsely suggest that the parental phenotypes are encoded by genes at different loci. Further characterization of these chromosomes and the physical locations of the *PDA* genes is needed to resolve this possibility. Despite these uncertainties, the current results indicate that the *PDA* genes belong to a unique family of cytochrome P450s but that there is structural diversity among these genes regardless of whether they are found at one or several loci.

Acknowledgements This work was supported in part by grant no. DCB-8517964 from NSF (H.D.V.), grant no. DE-FGO2-89ER14038 from DOE (H.D.V.), and National Research Service Award no. GM11009-03 from NIH (A.P.M.).

References

Ballance DJ (1991) Transformation systems for filamentous fungi and an overview of fungal gene structure. In: Leong SA, Berka RM (eds) Molecular industrial mycology – Systems and applications for filamentous fungi. Marcel Dekker, NY, pp 1–29

Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7:1513–1523

Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12:387–395

Gonzalez FJ (1988) The molecular biology of cytochrome P450s. Pharmacol Rev 40:243–288

Gonzalez-Candelas L, Kolattukudy PK (1992) Isolation and analysis of a novel inducible pectate lyase gene from the phytopathogenic fungus *Fusarium solani* f. sp. *pisi* (*Nectria haematococca*, mating population VI). J Bacteriol 174:6343–6349

Gotoh O (1982) An improved algorithm for matching biological sequences. J Mol Biol 162:705–708

Gotoh O, Fujii-Kuriyama Y (1989) Evolution, structure, and gene regulation of cytochrome P-450. In: Ruckpaul K, Rein H (eds) Basis and mechanisms of regulation of cytochrome P-450. Taylor and Francis, London, pp 196–245

Jacobson A (1987) Purification and fractionation of poly(A^+) RNA. Methods Enzymol 152:254–261

Kaergel E, Aoyama Y, Schunck W-H, Mueller H-G, Yoshida Y (1990) Comparative study on cytochrome P-450 of yeasts using specific antibodies to cytochromes $P\text{-}450_{alk}$ and $P\text{-}450_{14DM}$. Yeast 6:61–67

Kistler HC, VanEtten HD (1984a) Regulation of pisatin demethylase in *Nectria haematococca* and its influence on pisatin tolerance and virulence. J Gen Microbiol 130:2605–2613

Kistler HC, VanEtten HD (1984b) Three non-allelic genes for pisatin demethylation in the fungus *Nectria haematococca*. J Gen Microbiol 130:2595–2603

Kurtz S, Lundquist S (1984) Changing patterns of gene expression during sporulation in yeast. Proc Natl Acad Sci USA 81:7323–7327

Mackintosh SF, Matthews DE, VanEtten HD (1989) Two additional genes for pisatin demethylation and their relationship to the pathogenicity of *Nectria haematococca* on pea. Mol Plant-Microbe Interact 2:354–362

Matthews DE, VanEtten HD (1983) Detoxification of the phytoalexin pisatin by a fungal cytochrome P-450. Arch Biochem Biophys 224:494–505

Miao VPW, VanEtten HD (1992) Three genes for maackiain metabolism in the plant pathogen *Nectria haematococca*: meiotic instability and relationship to a new gene for pisatin demethylase. Appl Environ Microbiol 58:801–808

Miao VPW, Matthews DE, VanEtten HD (1991a) Identification and chromosomal locations of a family of cytochrome P-450 genes for pisatin detoxification in the fungus *Nectria haematococca*. Mol Gen Genet 226:214–223

Miao VPW, Covert SF, VanEtten HD (1991b) A fungal gene for antibiotic resistance on a dispensable ("B") chromosome. Science 254:1773–1776

Nebert DW, Nelson DR, Coon MJ, Estabrook RW, Feyereisen R, Fujii-Kuriyama Y, Gonzalez FJ, Guengerich FP, Gunsalus IC, Johnson EF, Loper JC, Sato R, Waterman MR, Waxman DJ (1991) The P450 superfamily: Update on new sequences, gene mapping, and recommended nomenclature. DNA Cell Biol 10:1–14

Needleman SB, Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol 48:443–453

Nelson DR, Strobel HW (1987) Evolution of cytochrome P-450 proteins. Mol Biol Evol 4:572–593

Nelson DR, Kamataki T, Waxman DJ, Guengerich FP, Estabrook RW, Feyereisen R, Gonzalez FJ, Coon MJ, Gunsalus IC, Gotoh O, Okuda K, Nebert DW (1993) The P450 superfamily: Update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. DNA Cell Biol 12:1–51

Porter TD, Coon MJ (1991) Cytochrome P-450: multiplicity of isoforms, substrates, and catalytic and regulatory mechanisms. J Biol Chem 266:13469–13472

Queen C, Korn LJ (1984) A comprehensive sequence analysis program for the IBM personal computer. Nucleic Acids Res 12:581–599

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. A laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Schäfer W, Straney D, Ciuffetti L, VanEtten HD, Yoder OC (1989) One enzyme makes a fungal pathogen, but not a saprophyte, virulent on a new host plant. Science 246:247–249

Soliday CL, Dickman MB, Kolattukudy PE (1989) Structure of the cutinase gene and detection of promoter activity in the 5'-flanking region by fungal transformation. J Bacteriol 171:1942–1951

Traynor PL, Levings CS (1986) Transcription of the S-2 maize mitochondrial plasmid. Plant Mol Biol 7:255–263

Turgeon GB, Garber RC, Yoder OC (1985) Transformation of the fungal maize pathogen *Cochliobolus heterostrophus* using the *Aspergillus nidulans amdS* gene. Mol Gen Genet 201:450–453

VanEtten HD, Matthews DE, Matthews P, Miao V, Maloney A, Straney D (1989) A family of genes for phytoalexin detoxification in the plant pathogen *Nectria haematoccoca*. In: Lugtenberg B (ed) Signal molecules in plants and plant-microbe interactions. (NATO ASI series, vol H36) Springer-Verlag, Berlin, pp 219–228

Weltring K-M, Turgeon BG, Yoder OC, VanEtten HD (1988) Isolation of a phytoalexin-detoxifying gene from the plant pathogenic fungus *Nectria haematococca* by detecting its expression in *Aspergillus nidulans*. Gene 68:335–344

Winter E, Yamamoto F, Almoguera C, Perucho M (1985) A method to detect and characterize point mutations in transcribed genes: Amplification and overexpression of the mutant c-Ki-*ras* allele in human tumor cells. Proc Natl Acad Sci USA 82:7575–7579