Antonie van Leeuwenhoek (2007) 91:179–189
DOI 10.1007/s10482-006-9109-3

ORIGINAL PAPER

The sequence of the isoepoxydon dehydrogenase gene of the patulin biosynthetic pathway in *Penicillium* species

Mary Ann Dombrink-Kurtzman

Received: 11 January 2006 / Accepted: 20 July 2006 / Published online: 17 October 2006

Abstract Interest in species of the genus *Penicillium* is related to their ability to produce the mycotoxin patulin and to cause spoilage of fruit products worldwide. The sequence of the isoepoxydon dehydrogenase (*idh*) gene, a gene in the patulin biosynthetic pathway, was determined for 28 strains representing 12 different *Penicillium* species known to produce the mycotoxin patulin. Isolates of *Penicillium carneum*, *Penicillium clavigerum*, *Penicillium concentricum*, *Penicillium coprobium*, *Penicillium dipodomyicola*, *Penicillium expansum*, *Penicillium gladioli*, *Penicillium glandicola*, *Penicillium griseofulvum*, *Penicillium paneum*, *Penicillium sclerotigenum* and *Penicillium vulpinum* were compared. Primer pairs for DNA amplification and sequencing were designed from the *P. griseofulvum idh* gene (GenBank AF006680). The two introns present were removed from the nucleotide sequences, which were translated to produce the IDH sequences of the 12 species for comparison. Phylogenetic relationships among the species were determined from rDNA (ITS1, 5.8 S, ITS2 and partial sequence of 28S rDNA) and from the *idh* nucleotide sequences minus the two introns. Maximum parsimony analysis showed trees based on rDNA and *idh* sequences to be congruent. It is anticipated that the genetic information obtained in the present study will aid in the design of probes, specific for patulin biosynthetic pathway genes, to identify the presence of these mycotoxigenic fungi.

Keywords Apple juice · Patulin · *Penicillium*

Introduction

The mycotoxin patulin is produced by 14 species of the genus *Penicillium* and is most likely to be found in fruits, silage and dung (Frisvad et al. 2004). Fungi capable of producing patulin can also be found in both soil and indoor air. In most cases, production of secondary metabolites (extrolites) is consistent from isolate to isolate in a species (Frisvad and Filtenborg 1983, 1989), suggesting that species-specific probes might be useful for establishing the level of patulin contamination. For example, all isolates of *Penicillium griseofulvum* examined produce patulin, griseofulvin, cyclopiazonic acid and roquefortine C. *Penicillium expansum*, the species most

M. A. Dombrink-Kurtzman (✉)
Mycotoxin Research Unit, National Center for Agricultural Utilization Research, Agricultural Research Service, United States Department of Agriculture, 1815 North University Street, Peoria, IL 61604-3902, USA
e-mail: dombrink@ncaur.usda.gov

commonly linked to patulin production and the species most often found on apples and other pomaceous fruits, also produces other toxins, such as roquefortine C and chaetoglobosin C, but these mycotoxins have not been found in rotting pomaceous fruit (Frisvad and Filtenborg 1989).

The most important species of *Penicillium* associated with food and described as patulin producers include, in addition to *P. expansum*, *P. griseofulvum*, *Penicillium carneum*, *Penicillium paneum* and *Penicillium sclerotigenum*. The foods associated with these latter four species are cereals and pasta (*P. griseofulvum*), beer, wine, rye-bread and dry meat products (*P. carneum*) (Frisvad and Filtenborg 1989), rye-bread (*P. paneum*) (Boysen et al. 1996) and *P. sclerotigenum* (yams) (Frisvad and Samson 2004); although patulin is not normally found in these foods, patulin is able to be produced when these fungi are grown on laboratory media. The remaining patulin-producing species can be isolated from soil, are primarily dung-associated and are not commonly found in foods, although they could be acquired if fruits, for example, apples, are dropped on the ground and are used in processing. These latter species include the following: *Penicillium clavigerum* (in soil, not food) (Svendsen and Frisvad 1994); *Penicillium concentricum* (rarely found in foods) (Leistner and Eckardt 1979); *Penicillium coprobium* (rarely found in foods) (Frisvad and Filtenborg 1989); *Penicillium dipodomyicola* (in rice, chicken feed, soil) (Frisvad et al. 1987); *Penicillium formosanum* (probably in dung); *Penicillium gladioli* (in *Gladiolus* bulbs) (Frisvad and Samson 2004); *Penicillium glandicola* (in silage, soil) (Frisvad and Filtenborg 1989); *Penicillium marinum* (in coastal sand); and *Penicillium vulpinum* (in dung, insects and soil) (Frisvad and Samson 2004). A compendium of information on species of *Penicillium* and the mycotoxins they produce has been described (Frisvad et al. 2004).

The patulin biosynthetic pathway contains at least ten enzymes in the direct pathway (Fedeshko 1992) and as many as a total of 19 enzymes when including those in side pathways. A number of biochemical studies to elucidate the patulin biosynthetic pathway have been performed (Sekiguchi and Gaucher 1978, 1979; Sekiguchi et al. 1983; Fedeshko 1992). The nucleotide sequence coding for the first pathway enzyme, 6-methylsalicylic acid synthase (6-MSAS), has been sequenced by Beck et al. (1990). The enzyme 6-MSAS is responsible for the condensing one molecule of malonyl-CoA and 3 molecules of acetyl-CoA; after extensive multi-step processing of 6-MSA, patulin is formed. Isoepoxydon dehydrogenase (IDH), the seventh enzyme in the pathway, is responsible for conversion of isoepoxydon to phyllostine in the terminal portion of the patulin biosynthetic pathway (Fedeshko 1992). *P. griseofulvum* strains had been used in the research of Beck et al. (1990) and Fedeshko (1992). Recently, both the full length *idh* gene and a fragment of the 6-*msas* gene have been cloned from *P. expansum* (White et al. 2006). Sequences of only these two genes (*6-msas* and *idh*) of the patulin pathway enzymes are present in GenBank.

In the present study, the sequence of the *idh* gene was determined for all of the species listed above, except for *P. formosanum* and *P. marinum*, to compare the extent of amino acid conservation at the site of catalytic activity. The phylogenetic relationships among the species were determined for both rDNA and *idh* gene sequences.

Materials and methods

Fungal isolates

Selection of patulin-producing species of *Penicillium* for determination of the nucleotide sequence of the *idh* gene was based on the mycotoxin determinations of Frisvad and Thrane (1987), as well as from suggestions by Jens C. Frisvad (personal communication). Twelve of the 14 patulin-producing species listed by Frisvad et al. (2004) and used in the study are listed in Table 1 and are maintained in the Agricultural Research Service Culture Collection (NRRL), Peoria, Illinois, USA.

DNA extraction

Cultures of different species of *Penicillium* were grown on PDA plates for approximately 14 days at 25°C. To each plate was added, 2–3 ml of 70%

Table 1 *Penicillium* species for which the isoepoxydon dehydrogenase gene sequence was determined

Penicillium strains	Strain designations[a]	
	NRRL	ATCC
Penicillium griseofulvum	5256	46037
Penicillium griseofulvum	2159A	24632
Penicillium griseofulvum	3523	36064
Penicillium griseofulvum	35258	549 (BFE)
Penicillium expansum	35231	
Penicillium expansum	6069	36200
Penicillium expansum	2304	28907
Penicillium expansum	32289	28879
Penicillium expansum	32293	28880
Penicillium expansum	35259	559 (BFE)
Penicillium clavigerum	1004	
Penicillium clavigerum	1003^T	10427^T
Penicillium carneum	25168^T	
Penicillium carneum	25170	
Penicillium concentricum	2034	
Penicillium coprobium	13626	64630
Penicillium dipodomyicola	35582	
Penicillium dipodomyicola	35583	
Penicillium gladioli	938	9437
Penicillium gladioli	939^T	10448^T
Penicillium glandicola	985	
Penicillium glandicola	2036^T	10450^T
Penicillium paneum	25159	
Penicillium paneum	25162	
Penicillium sclerotigenum	3461	
Penicillium sclerotigenum	22813	
Penicillium vulpinum	1002	9257
Penicillium vulpinum	2031^T	10426^T

[a] NRRL, ARS Culture Collection, National Center for Agricultural Utilization Research, Peoria, IL. ATCC, American Type Culture Collection, Manassas, VA. BFE, Bundesforschungsanstalt für Ernährung-Institut für Chemie und Biologie, Karlsruhe, Germany

ethanol and surface growth was brought into suspension using a sterile rod, bent at a 90° angle. The suspension was put into a 1.5 ml Eppendorf tube and centrifuged for 5 min to precipitate the cellular material. The supernatant was discarded and 375 µl of DNA extraction buffer (200 mM Tris, 250 mM NaCl, 25 mM EDTA, pH 8.5, 0.5% SDS) and 125 µl of glass beads (dia. 0.5 mm) (Scientific Industries Inc., Bohemia, NY) were added to a pellet. The tube was vortexed on the TurboMix (Scientific Industries Inc.) for 5–10 min; 350 µl of 2×CTAB (hexadecyltrimethylammonium bromide) buffer was added to the tube of broken cells, which was vortexed again for approximately 30 s, followed by addition of 350 µl of chloroform. The tube was carefully vortexed and then microfuged for 10 min at maximum speed to separate the emulsion. The upper aqueous phase was carefully removed, placed into a new 1.5 ml Eppendorf microfuge tube and 500 µl of isopropanol (– 20°C) was added to precipitate the DNA. The tube was centrifuged for 5 min and the supernatant was discarded. The DNA pellet was washed with 1 mL of 70% ethanol, centrifuged for 3 min and the supernatant was discarded. The DNA pellet was then dissolved in 100 µl of TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0). A 1:10 dilution of stock DNA in TE/10 buffer was the working stock of DNA.

PCR amplification

Before beginning PCR, concentrations of DNA and primers were determined and verified by agarose gel electrophoresis. Amplification of ITS and 28S rDNA domains D1/D2 was done using primers previously described (Peterson 2000). Primers used for determining the *idh* sequences are shown in Table 2. Symmetrical amplification was performed in a 96-well plate. The PCR mixture per well contained 6.3 µl of dH_2O, 4 µl of 10 × PCR buffer, 7.2 µl of deoxynucleotide (dNTP) mix (1.25 mM), 1.0 µl of external 5′ end primer (10 pmol/µl), 1.0 µl of external 3′ end primer (10 pmol/µl), 0.4 µl of *Taq* Polymerase (Sigma Chemical Company, St. Louis, MO) and 20 µl of genomic DNA in TE/10 buffer. The plate was covered with a rubber mat, vortexed, and spun briefly in a centrifuge to collect the liquid in the bottom of the wells. The plate was put into a PTC-100^{TM} Programmable Thermal Controller (MJ Research, Inc., Waltham, MA) and run using the following program: 35 × (94°C, 1 min; 50°C, 55 s; 72°C, 2 min), followed by 94°C, 1 min; 50°C, 55 s; 72°C, 10 min (final extension step). After amplification, all PCR products were visualized by agarose gel electrophoresis. If bands of varying sizes were observed, agarose gel purification of the appropriate sized bands was performed. If a single band of the appropriate size was observed, the PCR products were transferred to a MultiScreen plate (MANU03050, Millipore). Approximately 50 µl of dH_20 were added to each well and PCR products were dried under vacuum. An

Table 2 Oligonucleotide primers used in this study for sequencing the isoepoxydon dehydrogenase gene

Primer	Sequence (5′ → 3′)	Position
IDH 2076F	GCCCATGTGCTCATTACAGG	2067–2086
IDH 2667R	TGGGACAATTCCTGAACATGC	2647–2667
IDH 2195F	CAATGTGTCGTACTGTGCCC	2195–2214
IDH 2793R	ACCTTCAGTCGCTGTTCCTC	2774–2793
IDH 2050F	CAACTGGTCTCAAAGGTGCC	2050–2069
IDH 2887R	CAACGTGAATTCCGCCATCAACCAAC	2862–2887
IDH 1909F	GGTCTAAGAGAGCCGATTGTCTC	1909–1931
P.e. IDH 1F	ATGGTGCTCATCACTGGTC	1–19
IDH 2016F	CAACTTTCAAATTTATTCGC	2016–2035
IDH 2933R	CTCATGGCTTATTCTGAGTTC	2914–2933

Primer pairs specific for the *P. griseofulvum* (formerly known as *P. urticae*) isoepoxydon dehydrogenase (*idh*) gene (GenBank AF006680) that were used for sequencing. All sets of primers were synthesized by Sigma-Genosys (The Woodlands, TX). Primers used as pairs were: IDH 2067F and IDH 2667R; IDH 2016F and IDH 2793R; IDH 2195F and IDH 2793R; IDH 2195F and IDH 2933R; IDH 2050F and IDH 2887R; IDH 2067F and IDH 2887R; and IDH 1909F and IDH 2887R. The latter primer pair was used only to obtain sequences for strains of *P. griseofulvum*. Primer pair IDH 2195F and IDH 2793R had previously been used to compare isolates from apple orchards in the United Kingdom (Paterson et al. 2000, 2003). The sequence for one primer (*P.e.* IDH 1F) was based on the first 19 nucleotides of the *P. expansum idh* gene (GenBank DQ084388) which was used paired with IDH 2887R

additional wash of 100 µl of dH_20 and drying followed. Purified DNA was solubilized in 50 µl of dH_20 and the plate was placed on a shaker for 20 min. DNA was then transferred to an unused PCR plate, which was covered with a foil seal and placed in the freezer until cycle sequencing.

Nucleotide sequencing

Sequencing of both strands of PCR products was performed using an Applied Biosystems MicroAmp 96-well plate and the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA). To each well was added: 3 µl of dH_2O, 2 µl of BigDye, 2 µl of Primer (1 pmol/µl) and 3 µl of DNA. The plate was covered with a wax film, vortexed and spun. The thermal cycler program had 35 cycles, consisting of a denaturation step at 94°C for 30 s, an annealing step at 50°C for 15 s, and an elongation step at 60°C for 4 min. The cycling reaction product was purified and prepared for sequencing. To each well was added: 34 µl of 100% ethanol and 4 µl of 10 M ammonium acetate; contents of the wells were mixed with a pipette tip. The plate was covered with a rubber mat and placed on a shaker for 20 min. The plate was centrifuged at 3,600 rpm for 45 min. The rubber mat was removed and the plate was inverted on a paper towel; the plate was then pulse spun. To each well was added 150 µl of 75% ethanol, the rubber mat was replaced and the plate was centrifuged for 10 min. After centrifuging, the plate was allowed to drain over a paper towel and then pulse spun. The plate was dried in a swinging bucket vacuum centrifuge for 20–30 min. Ten microliters of formamide (Hi-Di) were added to each well for sequencing. Samples were sequenced on an ABI 3730 DNA Analyzer or ABI 3100 DNA Analyzer. Sequence analysis was conducted using Sequencher software v4.2 (Gene Codes Corporation).

Sequence analysis

Phylogenetic relationships were determined from rDNA sequences and *idh* sequences using the maximum parsimony (MP) program of PAUP* 4.0 (Swofford, 1998) with heuristic searches. Further comparisons were made with the neighbor-joining (NJ) program included in PAUP* 4.0 using the Kimura 2-parameter distance correction. Confidence limits for phylogenetic trees were estimated from bootstrap analyses (1,000 replications).

Results

The nucleotide sequences of the *idh* gene (either full length or partial *idh* sequence) for all 28

strains of the *Penicillium* species were obtained and translated to determine the amino acid sequences. Following removal of the two introns present in the *idh* sequence, amino acid sequences were compared (Fig. 1). Additionally, ribosomal DNA (ITS1, 5.8 S, ITS2 and partial sequence of

```
Pgri 2159A  MVLTTGLKGA  HVLITGGTRG  MGEAMVHKFL  EEEANVSYCA  RTVTNKEFDE  FHATLPEGNT
Pgri  3523  ..........  ..........  ..........  ..........  ..........  ..........
Pgri  5256  ..........  ..........  ..........  ..........  ..........  ..........
Pgri 35258  ..........  ..........  ..........  ..........  ..........  ..........
Pdip 35582  ..........  ..........  ..........  ..........  ..........  ..........
Pdip 35583  ..........  ..........  ..........  ..........  ..........  ..........
Pexp  2304  ...       .  ..........  ..........  Q.........  .....T.Y.D  .YS..A....
Pexp  6069  ...       .  ..........  ..........  Q.........  .....T.Y.D  .YS..A....
Pexp 32289  .....     .  ..........  ..........  Q.........  .....T.Y.D  .YS..A....
Pexp 32293  .....     .  ..........  ..........  Q.........  .....T.Y.D  .YS..A....
Pexp 35231  ...       .  ..........  ..........  Q.........  .....T.Y.D  .YS..A....
Pexp 35259  ..        .  ..........  ..........  Q.........  .....T.Y.D  .YS..A....
Pcar 25168  ...S......   ..........  ..........  ..........  .....T.Y.D  .YP.......
Pcar 25170  ...S......   ..........  ..........  ..........  .....T.Y.D  .YP.......
Pcla  1003  ..........   .....       ..........  Q.........  .....T.Y.D  .YP..A....
Pcla  1004  ..........   ........    ..........  ..........  .....T.Y.D  .YL..A....
Pcon  2034  ...       .  ..........  ..........  Q.........  .....T....  .YP......I
Pcop 13626  ..........   ....        ..........  ..........  .....T....  .YS.......
Pgld   938  ......    .  ..........  ..........  Q.........  .....T.Y.D  .YP..A....
Pgld   939  ......    .  ..........  ..........  Q.........  .....T.Y.D  .YP..A....
Pgln   985  .         .  ..........  ..........  Q.........  .S...T...D  .YP..A....
Pgln  2036  .         .  ..........  ..........  Q.........  .S...T...D  .YP..AR...
Ppan 25159  ...       .  ..........  ......Q...  ..........  .....T.Y.D  .YP.......
Ppan 25162  ......    .  ..........  ......Q...  ..........  .....T.Y.D  .YP.......
Pscl  3461  .         .  ..........  ..........  Q.........  .....T.Y..  .YS..A....
Pscl 22813  ..        .  ..........  ..........  Q.........  .....T.Y..  .YS..A....
Pvul  1002  ......    .  ..........  ..........  Q.........  .....T.Y..  .YS.......
Pvul  2031  ......    .  ..........  ..........  Q.........  .....T.Y..  .YS.......

Pgri 2159A  ARAVGTAFDV  ASKDAIVQWV  ESSAERLGRI  DVIIANASPM  HMEGETEHWE  SSFAIDVMGF
Pgri  3523  ..........  ..........  ..........  ..........  ..........  ..........
Pgri  5256  ..........  ..........  ..........  ..........  ..........  ..........
Pgri 35258  ..........  ..........  ..........  ..........  ..........  ..........
Pdip 35582  ..........  ..........  ..........  ..........  ..........  ..........
Pdip 35583  ..........  ..........  ..........  ..........  ..........  ..........
Pexp  2304  ..........  ....SL.K..  ..........  ..........  ..........  ..........
Pexp  6069  ..........  ....SL.K..  ..........  ..........  ..........  ..........
Pexp 32289  ..........  ....SL.K..  ..........  ..........  ..........  ..........
Pexp 32293  ..........  ....SL.K..  ..........  ..........  ..........  ..........
Pexp 35231  ..........  ....SL.K..  ..........  ..........  ..........  ..........
Pexp 35259  ..........  ....SL.K..  ..........  ..........  ..........  ..........
Pcar 25168  ..........  .....L....  ....Q.....  ..........  ..........  ..........
Pcar 25170  ..........  .....L....  ....Q.....  ..........  ..........  ..........
Pcla  1003  ..........  .....L.K..  ..........  ..........  ..........  ..........
Pcla  1004  ..........  .....L.K..  ..........  ..........  ..........  ..........
Pcon  2034  ..........  ..........  ..........  ..........  ..........  ..........
Pcop 13626  ..........  ..........  ..........  ..........  ..........  ..........
Pgld   938  ..........  .....L.K..  ..........  ..........  ..........  ..........
Pgld   939  ..........  .....L.K..  ..........  ..........  ..........  ..........
Pgln   985  ..........  ..........  .R........  ..........  ..........  ..........
Pgln  2036  ..........  ..........  .R........  ..........  ..........  ..........
Ppan 25159  ..........  .....L....  ....Q.....  ..........  ......G...  ..........
Ppan 25162  ..........  .....L....  ....Q.....  ..........  ......G...  ..........
Pscl  3461  ..........  .....L.K..  ..........  ..........  ..........  ..........
Pscl 22813  ..........  .....L.K..  ..........  ..........  ..........  ..........
Pvul  1002  ..........  .....L....  ..........  ..........  ..........  ..........
Pvul  2031  ..........  .....L....  ..........  ..........  ..........  ..........
```

Fig. 1 Multiple sequence alignment of deduced amino acid sequences of the *idh* sequences from *P. griseofulvum* (*Pgri*), *P. expansum* (*Pexp*), *P. clavigerum* (*Pcla*), *P. carneum* (*Pcar*), *P. concentricum* (*Pcon*), *P. coprobium* (*Pcop*), *P. dipodomyicola* (*Pdip*), *P. gladioli* (*Pgld*), *P. glandicola* (*Pgln*), *P. paneum* (*Ppan*), *P. sclerotigenum* (*Pscl*) and *P. vulpinum* (*Pvul*). Identical sequences are indicated with a dot, except where indicated by an alternative amino acid residue. Partial and complete isoepoxydon dehydrogenase sequences are given. Conserved amino acids (Fedeshko 1992) are shown in bold. Underlined residues are basic amino acids and are found only in $NADP^+$ specific dehydrogenases (Scrutton et al. 1990)

```
Pgri 2159A  VELVRAATPY  LEKSPQASII  VQSSFMGREF  YRSPPAAYGP  CKAAQLQHVQ  ELSHFLGPKG
Pgri  3523  ..........  ..........  ..........  ..........  ..........  ..........
Pgri  5256  ..........  ..........  ..........  ..........  ..........  ..........
Pgri 35258  ..........  ..........  ..........  ..........  ..........  ..........
Pdip 35582  ..........  ..........  ..........  ..........  ..........  ..........
Pdip 35583  ..........  ..........  ..........  ..........  ..........  ..........
Pexp  2304  ....K.....  ..........  ..........  ..........  ..........  ..........
Pexp  6069  ....K.....  ..........  ..........  ..........  ..........  ..........
Pexp 32289  ....K.....  ..........  ..........  ..........  ..........  ..........
Pexp 32293  ....K.....  ..........  ..........  ..........  ..........  ..........
Pexp 35231  ....K.....  ..........  ..........  ..........  ..........  ..........
Pexp 35259  ....K.....  ..........  ..........  ..........  ..........  ..........
Pcar 25168  ....K.....  ..........  ..........  ..........  ..........  ..........
Pcar 25170  ....K.....  ..........  ..........  ..........  ..........  ..........
Pcla  1003  ....K.....  ..........  ..........  ..........  ..........  ..........
Pcla  1004  ....K.....  ..........  ..........  ..........  ..........  ..........
Pcon  2034  ....K.....  ....T.....  ..........  ..........  ..........  ..........
Pcop 13626  ..........  ....S.....  ..........  ..........  ..........  ..........
Pgld   938  ....K.....  ..........  ..........  ..........  ..........  ..........
Pgld   939  ....K.....  ..........  ..........  ..........  ..........  ..........
Pgln   985  ....K.....  ....S.....  ..........  ..........  ..........  ..........
Pgln  2036  ....K.....  ..N.......  .......Q..  ..........  ..........  ......E.R.
Ppan 25159  ....K.....  ..........  ..........  ..........  ..........  ..........
Ppan 25162  ....K.....  ..........  ..........  ..........  ..........  ..........
Pscl  3461  ....K.....  ..........  ..........  ..........  ..........  ..........
Pscl 22813  ....K.....  ..........  ..........  ..........  ..........  ..........
Pvul  1002  ..........  ..........  ..........  ..........  ..........  ..........
Pvul  2031  ..........  ..........  ..........  ..........  ..........  ..........

Pgri 2159A  IRVNAISPGP  ILCKGGPWEL  YSKINPEWVE  EQRLKVPLKR  LGGPTEVAAV  AVFLASPLAS
Pgri  3523  ..........  ..........  ..........  ..........  ..........  ..........
Pgri  5256  ..........  ..........  ..........  ..........  ..........  ..........
Pgri 35258  ..........  ..........  ..........  ..........  ..........  ..........
Pdip 35582  ..........  ..........  ..........  ..........  ..........  ..........
Pdip 35583  ..........  ..........  ..........  ..........  ..........  ..........
Pexp  2304  ..........  V.........  ..........  .....I....  .........N.  ..........
Pexp  6069  ..........  V.........  ..........  .....I....  .........N.  ..........
Pexp 32289  ..........  V.........  ..........  .....I....  .........N.  ..........
Pexp 32293  ..........  V.........  ..........  .....I....  .........N.  ..........
Pexp 35231  ..........  V.........  ..........  .....I....  .........N.  ..........
Pexp 35259  ..........  V.........  ..........  .....I....  .........N.  ..........
Pcar 25168  ..........  ..........  ..........  .....I....  ....K.....  ..........
Pcar 25170  ..........  ..........  ..........  .....I....  ....K.....  ..........
Pcla  1003  ..........  ..........  ...T......  .....I....  .........S.  ........S.
Pcla  1004  ..........  ..........  ...T......  .....I....  .........S.  ........S.
Pcon  2034  ..........  ........D.  ......D...  .....I....  ..........  ..........
Pcop 13626  ..........  ..........  ..........  .....I....  ..........  ..........
Pgld   938  ..........  ..........  ...T......  ...V.I....  ..........  ..........
Pgld   939  ..........  ..........  ...T......  ...V.I....  ..........  ..........
Pgln   985  ..........  ..........  ...T......  .....I....  ..........  ..........
Pgln  2036  .W........  ..........  ...T......  .....I....  ..........  ..........
Ppan 25159  ..........  ..........  ..........  .....I....  ....K.....  ....S.....
Ppan 25162  ..........  ..........  ..........  .....I....  ....K.....  ....S.....
Pscl  3461  ..........  ..........  ...T......  .....I....  ..........  ..........
Pscl 22813  ..........  ..........  ...T......  .....I....  ..........  ..........
Pvul  1002  ..........  ..........  ...T......  .....I....  .........D.  ..........
Pvul  2031  ..........  ..........  ...T......  .....I....  .........D.  ..........
```

Fig. 1 continued

28S rDNA) for all 28 isolates was also sequenced and compared to verify that species had been correctly identified. The *idh* sequence of *P. griseofulvum* NRRL 2159A was used as a reference because it had been the first *idh* sequence deposited in GenBank (accession AF006680). It was also used to align the *idh* sequences of the *Penicillium* species and to determine the location of the two introns (present at nucleotides 2,086–2,139 and 2,383–2,435 for the *idh* sequence in GenBank). In Fig. 1, the nine conserved amino acids in the IDH sequence described by Fedeshko (1992) are indicated in bold; eight of the nine conserved amino acids

```
Pgri 2159A  FVSGTNMLVD  GGIHVGTQF.
Pgri  3523  ..........  ..........
Pgri  5256  ..........  ..........
Pgri 35258  ..........  ..........
Pdip 35582  ..........  ..........
Pdip 35583  ..........  ..........
Pexp  2304  ...
Pexp  6069  .....
Pexp 32289  .........
Pexp 32293  ..........
Pexp 35231  ..........
Pexp 35259  ..
Pcar 25168  ..........  .....
Pcar 25170  ..........  .....
Pcla  1003  ..T...
Pcla  1004  ..T..
Pcon  2034  ..........  ..........
Pcop 13626  ..........  ....
Pgld   938  ..T.......
Pgld   939  ..T.......
Pgln   985  ..........  .....
Pgln  2036  ..........  .....
Ppan 25159  .L....V...  .....
Ppan 25162  ......V...
Pscl  3461  ..T...V...  ..........
Pscl 22813  ..T...V...  ..........
Pvul  1002  ......L...  .....
Pvul  2031  ......L...  ..........
```

Fig. 1 continued

were identical in all of the *Penicillium* species examined in this study. The ninth conserved amino acid, lysine (a basic amino acid) at position 46, was present in all strains of *P. griseofulvum* and *P. dipodomyicola* examined, but lysine was replaced by threonine (a polar amino acid) in all of the other *Penicillium* species. Amino acids present at other positions for many of the *Penicillium* species had common substitutions: glutamine substituted for glutamic acid (position 31); tyrosine substituted for phenylalanine (position 48); aspartic acid substituted for glutamic acid (position 50); tyrosine substituted for histidine (position 52); serine or proline substituted for alanine (position 53); alanine substituted for proline (position 56); leucine substituted for isoleucine (position 86); lysine substituted for glutamine (position 88); lysine substituted for arginine (position 125); threonine substituted for isoleucine (position 204); and isoleucine substituted for valine (position 216). Additionally, there were certain substitutions not present in any of the other *Penicillium* species that were shared by all of the strains of *P. expansum*; this included serine substituted for alanine (position 85), valine substituted for isoleucine (position 191), and asparagine substituted for alanine (position 229) (Fig. 1).

Phylogenetic trees showing placement of strains of the *Penicillium* species are shown for the *idh* sequences (Fig. 2) and for the rDNA sequences (Fig. 3). The nucleotide sequences of the *idh* gene and the rDNA genes of *Byssochlamys nivea* (NRRL 32565^{T}) (Dombrink-Kurtzman and Engberg 2006) were used as outliers for the respective trees. *B. nivea* was chosen as an outlier because it is capable of producing patulin and it is not a member of *Penicillium* subgenus *Penicillium*. When two or more strains of a given species were analyzed, the strains usually had identical nucleotide sequences. Species relationships on phylogenetic trees determined from the *idh* sequences (Fig. 2) and from the rDNA sequences (Fig. 3) were similar on branches with strong bootstrap support. In addition, results were nearly identical whether analyzed from maximum parsimony or neighbor-joining analysis (data not shown).

Discussion

The present study was focused on *Penicillium* species capable of producing the mycotoxin patulin. Research by Frisvad and collaborators was used as a resource for information on the production of secondary metabolites (extrolites) by species of *Penicillium* and has been essential for understanding which species produce patulin (Frisvad and Thrane 1987; Frisvad et al. 1987; Frisvad and Filtenborg 1989; Frisvad and Sampson 2004). Patulin production, determined under laboratory conditions, ranged from 65.72 to 294.05 μg/ml for strains of *P. expansum* and from 812.3 to 1,927.65 μg/ml for strains of *P. griseofulvum* (Dombrink-Kurtzman and Blackburn 2005). Surprisingly, *P. griseofulvum* produced approximately 10-fold more patulin than strains of *P. expansum*, the species most cited as the primary fungus associated with patulin contamination of fruits, primarily apples (Dombrink-Kurtzman and Blackburn 2005).

Comparison of the amino acids sequences of all *P. griseofulvum* and *P. dipodomyicola* strains examined in the present study revealed them to

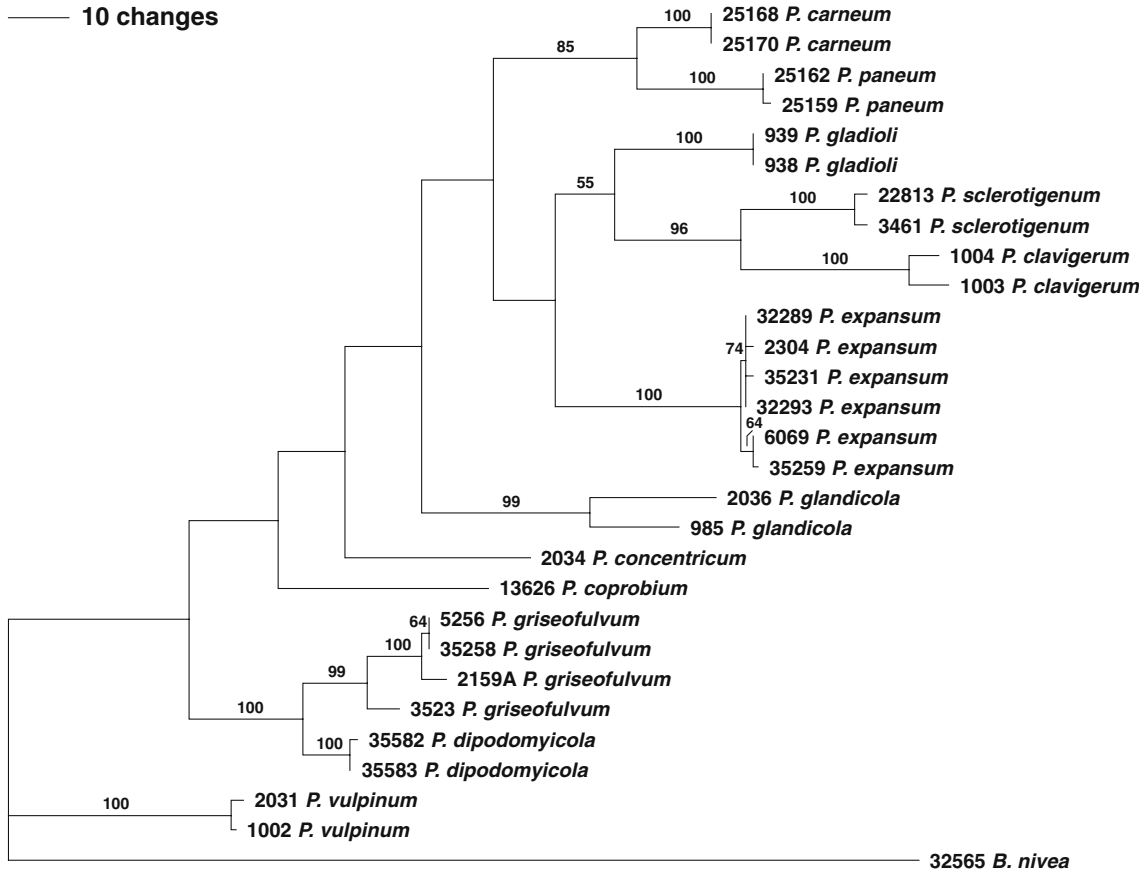

Fig. 2 Phylogenetic tree showing placement of *Penicillium* species as represented by 1 of 4 most parsimonious trees derived from maximum parsimony analysis of the *idh* gene with introns 1 and 2 removed. Tree length = 683, consistency index (CI) = 0.638, retention index (RI) = 0.786, rescaled consistency index (RC) = 0.502, homoplasy index (HI) = 0.362. Branch lengths are indicated by the marker bar, and numbers at nodes are bootstrap values determined from 1,000 replications. Bootstrap values below 50% are not displayed. *Byssochlamys nivea* NRRL 32565^{T} was chosen as the outgroup species to root the tree

be identical (Fig. 1), although some nucleotide differences existed at the third position coding for an amino acid. *P. griseofulvum* and *P. dipodomyicola* are closely related and are commonly found on dry cereals and seeds (Frisvad et al. 2004). *P. griseofulvum* and *P. dipodomyicola* were also the only *Penicillium* species examined to have the conserved amino acid lysine present at position 46 in the IDH protein. In all of the other species, the conserved lysine is replaced by threonine. The strains having lysine at position 46 produced the most patulin of the strains examined. The significance of this amino acid substitution relates to the inability of the enzyme IDH to function optimally if the conserved amino acid (lysine) is not present to participate in binding the cofactor $NADP^{+}$. The conserved basic amino acid lysine is thought to be involved in stabilizing the extra negative phosphate charge on the $NADP^{+}$ molecule in dehydrogenases requiring the cofactor $NADP^{+}$ as a reducing coenzyme (Scrutton et al. 1990). Replacement of a conserved amino acid by another amino acid can have different effects in the resulting protein (isoepoxydon dehydrogenase), depending upon whether the substitution involves a similar or different type of amino acid. If the type of amino acid substitution is radically different, the conformation of the protein may be affected, leading to a perturbation of the active site of the enzyme and possibly lower functionality of the resulting protein.

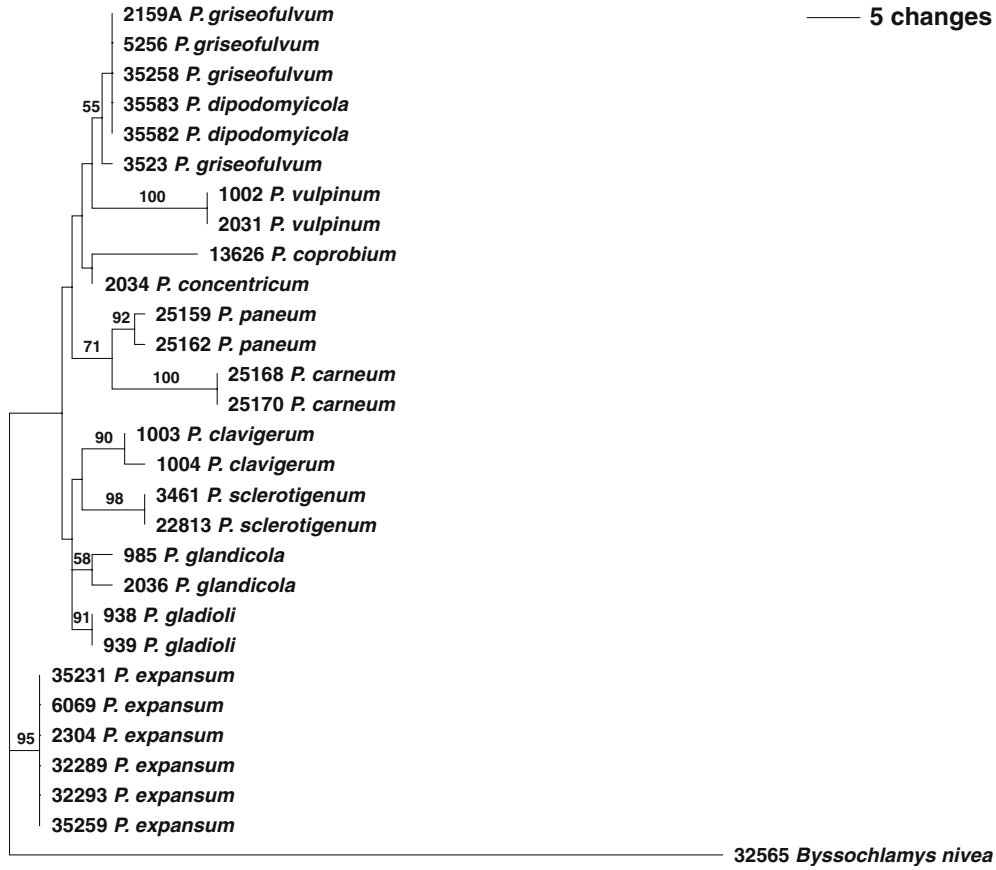

Fig. 3 Phylogenetic tree showing placement of *Penicillium* species as represented by 1 of 18 most parsimonious trees derived from maximum parsimony analysis of rDNA (ITS1, 5.8 S, ITS2 and partial sequences of 28S rDNA). Tree length = 157, consistency index (CI) = 0.815, retention index (RI) = 0.809, rescaled consistency index (RC) = 0.660, homoplasy index (HI) = 0.185. Branch lengths are indicated by the marker bar, and numbers at nodes are bootstrap values determined from 1,000 replications. Bootstrap values below 50% are not displayed. *Byssochlamys nivea* NRRL 32565^{T} was chosen as the outgroup species to root the tree

Alignments of introns 1 and 2 were done for all 28 of the *Penicillium* strains; phylogenetic trees were determined for introns 1 and 2 (data not shown) and analyzed by maximum parsimony. The trees of introns 1 and 2 were congruent with trees based on the *idh* nucleotide sequences (Fig. 2), suggesting that the introns and exons had evolved in tandem. Intron 1 contained 50–69 nucleotides. The largest number of nucleotides was present in *P. coprobium*, due to the presence of a stretch of adenine nucleotides. Intron 2 contained nucleotides ranging from 49 to 54.

Interestingly, the respective nucleotide sequence for introns 1 and 2 (*idh* nucleotide sequences for all strains have been deposited in GenBank) were highly conserved for strains of a particular species. When analyzed by maximum parsimony, the *idh* (Fig. 2) and rDNA (Fig. 3) sequences were also congruent. *P. griseofulvum* and *P. dipodomyicola*, which are both members of the series *Urticicolae*, were found to be grouped together for both the *idh* and rDNA sequences. This would suggest that for the different *Penicillium* species examined in this study, lateral gene transfer had not occurred among the different strains.

P. carneum and *P. paneum* can occur in very acidic products, such as silage (Frisvad and Samson 2004). Results of the maximum parsimony analyses in the present study indicated that *P. carneum* and *P. paneum* were closely related (Figs. 2 and 3). Although *P. carneum* and

P. paneum have been reported to be indistinguishable using traditional morphological and physiological characteristics, they have a 12 nucleotide difference in the ITS regions (Boysen et al. 1996); results, such as these, have led to the conclusion that molecular characterization is essential for correct identification of fungal strains.

It is anticipated that the information gained from this study of *idh* sequences of *Penicillium* species with the ability to produce patulin will be beneficial for the design and use of molecular probes for identification of these mycotoxigenic fungi. Work in progress has identified unique species-specific sequences present in intron 2 for the design of 5′ primers. When utilized with specific 3′ primers, this will allow for generation of species-specific amplicons for rapid, accurate identification of patulin-producing species of *Penicillium*.

Nucleotide sequence accession numbers

The nucleotide sequences for the *idh* sequences determined in this study have been deposited with GenBank and have been assigned the following accession numbers: *P. carneum* NRRL 25168 (DQ343637) and NRRL 25170 (DQ343638); *P. clavigerum* NRRL 1003 (AY885572) and NRRL 1004 (AY885571); *P. concentricum* NRRL 2034 (DQ343630); *P. coprobium* NRRL 13626 (DQ343633); *P. dipodomyicola* NRRL 35582 (DQ343643) and NRRL 35583 (DQ343644); *P. expansum* NRRL 32293 (DQ343640), NRRL 32289 (DQ343639) and NRRL 35259 (DQ343642); *P. gladioli* NRRL 938 (DQ343625) and NRRL 939 (DQ343626); *P. glandicola* NRRL 985 (DQ343627) and NRRL 2036 (DQ343631) (an extra guanine at position 83 was believed to be a sequencing artifact and was not included in the submission to GenBank); *P. griseofulvum* NRRL 35258 (DQ343641); *P. paneum* NRRL 25159 (DQ343635) and NRRL 25162 (DQ343636); *P. sclerotigenum* NRRL 3461 (DQ343632) and NRRL 22813 (DQ343634); and *P. vulpinum* NRRL 1002 (DQ343628) and NRRL 2031 (DQ343629). Nucleotide sequences for the *idh* gene that were previously deposited in GenBank and included here for comparisons are *P. griseofulvum* NRRL 3523 (AY885565.1), NRRL 5256 (AY885566.1) and NRRL 2159A (AY885567.1); and *P. expansum* NRRL 2304 (AY885568), NRRL 6069 (AY885570) and NRRL 35231 (AY885569.1) (Dombrink-Kurtzman 2006).

The nucleotide sequences for the ITS and 28S rDNA domains have been deposited in GenBank and have been assigned the following GenBank numbers: *P. carneum* NRRL 25168 (DQ339564) and NRRL 25170 (DQ339566); *P. clavigerum* NRRL 1003 (DQ339555) and NRRL 1004 (DQ339560); *P. concentricum* NRRL 2034 (DQ339561); *P. coprobium* NRRL 13626 (DQ339559); *P. dipodomyicola* NRRL 35582 (DQ339550) and NRRL 35583 (DQ339570); *P. expansum* NRRL 2304 (DQ339556), NRRL 6069 (DQ339562), NRRL 32289 (DQ339547), NRRL 32293 (DQ339552), NRRL 35231 (DQ339558) and NR RL 35259 (DQ339548); *P. gladioli* NRRL 938 (DQ339563) and NRRL 939 (DQ339568); *P. glandicola* NRRL 985 (DQ339573) and NRRL 2036 (DQ339565); *P. griseofulvum* NRRL 3523 (DQ339549), NRRL 5256 (DQ339557), NRRL 2159A (DQ339551) and NRRL 35258 (DQ339553); *P. paneum* NRRL 25159 (DQ339554) and NRRL 25162 (DQ339571); *P. sclerotigenum* NRRL 3461 (DQ339567) and NRRL 22813 (DQ339574); and *P. vulpinum* NRRL 1002 (DQ339572) and NRRL 2031 (DQ339569).

Acknowledgements Special thanks are given to Amy E. Engberg for excellent technical assistance and to Dr. Cletus P. Kurtzman for generation of phylogenetic trees. Names of equipment and chemical supplies are necessary to report factually on experimental methods; however, the USDA neither guarantees nor warrants the standard of the products, and the use of the name by USDA implies no approval of the products to the exclusion of others that may also be suitable.

References

Beck J, Ripka S, Siegner A, Schiltz E, Schweizer E (1990) The multifunctional 6-methylsalicylic acid synthase gene of *Penicillium patulum*. Its gene structure relative to that of other polyketide synthases. Eur J Biochem 192:487–498

Boysen M, Skouboe P, Frisvad JC, Rossen L (1996) Reclassification of the *Penicillium roqueforti* group into three species on the basis of molecular genetic and biochemical profiles. Microbiology 142:541–549

Dombrink-Kurtzman MA (2006) The isoepoxydon dehydrogenase gene of the patulin metabolic pathway differs for *Penicillium griseofulvum* and *Penicillium expansum*. Anton van Leeuw 89:1–8

Dombrink-Kurtzman MA, Blackburn JA (2005) Evaluation of several cultural media for production of patulin by *Penicillium species*. Int J Food Microbiol 98:241–248

Dombrink-Kurtzman MA, Engberg AE (2006) *Byssochlamys nivea* with patulin producing capability has an isoepoxydon dehydrogenase gene (*idh*) with sequence homology to *Penicillium expansum* and *P. griseofulvum*. Mycol Res 110:1111–1118

Fedeshko RW (1992) Polyketide enzymes and genes in *Penicillium urticae*. Ph.D. Dissertation, University of Calgary, Calgary, Alberta, Canada

Frisvad JC, Filtenborg O (1983) Classification of terverticillate penicillia based on profiles of mycotoxins and other secondary metabolites. Appl Environ Microbiol 46:1301–1310

Frisvad JC, Filtenborg O (1989) Terverticillate penicillia: chemotaxonomy and mycotoxin production. Mycologia 81:837–861

Frisvad JC, Samson RA (2004) Polyphasic taxonomy of *Penicillium* subspecies *Penicillium*. A guide to identification of food and air-borne terverticillate Penicillia and their mycotoxins. Stud Mycol 49:1–174

Frisvad JC, Thrane U (1987) Standardized high-performance liquid chromatography of 182 mycotoxins and other fungal metabolites based on alkylphenone retention indices and UV–VIS spectra (diode array detection). J Chromatogr 404:195–214

Frisvad JC, Filtenborg O, Wicklow DT (1987) Terverticillate Penicillia isolated from underground seed caches and cheek pouches of banner-tailed kangaroo rats (*Dipodomys spectabilis*). Can J Bot 65:765–773

Frisvad JC, Smedsgaard J, Larsen TO, Samson RA (2004) Mycotoxins, drugs and other extrolites produced by species in *Penicillium* subgenus *Penicillium*. Stud Mycol 49:201–242

Leistner L, Eckardt C (1979) Vorkommen toxinogener Penicillien bei Fleischerzeugnissen. Fleischwirtschaft 59:1892–1896

Paterson RRM, Archer S, Kozakiewicz Z, Lea A, Locke T, O'Grady E (2000) A gene probe for the patulin metabolic pathway with potential for use in patulin and novel disease control. Biocontrol Sci Technol 10:509–512

Paterson RRM, Kozakiewicz Z, Locke T, Brayford D, Jones SCB (2003) Novel use of the isoepoxydon dehydrogenase gene probe of the patulin metabolic pathway and chromatography to test penicillia isolated from apple production systems for the potential to contaminate apple juice with patulin. Food Microbiol 20:359–364

Peterson SW (2000) Phylogenetic analysis of *Penicillium* species based on ITS and LSU-rDNA nucleotide sequences. In: Samson RA, Pitt JI (eds) Integration of modern taxonomic methods for *Penicillium* and *Aspergillus* classification. Harwood Academic Publishers, Amsterdam, pp 163–178

Scrutton NS, Berry A, Perham RN (1990) Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. Nature 343:38–43

Sekiguchi J, Gaucher GM (1978) Identification of phyllostine as an intermediate of the patulin pathway in *Penicillium urticae*. Biochemistry 19:1785–1791

Sekiguchi J, Gaucher GM (1979) Isoepoxydon, a new metabolite of the patulin pathway in *Penicillium urticae*. Biochem J 182:445–453

Sekiguchi J, Shimamoto T, Yamada Y, Gaucher GM (1983) Patulin biosynthesis: enzymatic and nonenzymatic transformations of the mycotoxin (E)-ascladiol. Appl Environ Microbiol 45:1939–1942

Svendsen A, Frisvad JC (1994) A chemotaxonomic study of the terverticillate Penicillia based on high performance liquid chromatography of secondary metabolites. Mycol Res 98:1317–1328

Swofford DL (1998) PAUP* 4.0: phylogenetic analysis using parsimony (and other methods). Sinauer Associates, Sunderland, MA, USA

White S, O' Callaghan J, Dobson ADW (2006) Cloning and molecular characterization of *Penicillium expansum* genes upregulated under conditions permissive for patulin biosynthesis. FEMS Microbiol Lett 255:17–26