Structural organization of polygalacturonase-encoding genes from *Penicillium griseoroseum*

Andréa O. B. Ribon, Marisa V. Queiroz and Elza F. de Araújo

Departamento de Microbiologia / BIOAGRO, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil.

Abstract

The pectinolytic system of *Penicillium griseoroseum* has been studied as a model to investigate aspects of gene organization in filamentous fungi. Here we show that the endopolygalacturonase-coding genes previously isolated exist as single copies in the fungus genome. DNA blot analysis revealed the presence of corresponding genes in other *Penicillium* species, although only one or two genes were found in opposition to the endoPG gene family reported for other filamentous fungi. The nucleotide and amino acid sequences of *Penicillium* PG genes of retrieved from data banks were compared for intron length and number, codon usage, and consensus sequences for translation initiation sites. The introns are conserved in the same position, although there was no conservation of their nucleotide sequences. Other sequence features resemble those seen in *Aspergillus* and *Neurospora* genes.

Key words: *Penicillium griseoroseum*, filamentous fungi, pectinase, polygalacturonase, gene organization.

Received: August 26, 2002; accepted: November 14, 2002.

Introduction

The genus *Penicillium* genus is worldwide known for the production of secondary metabolites and extracellular enzymes of commercial value which include the pectinases, employed in the fruit juice industry during the stage of maceration of the pulp, juice liquefaction or depectinization (Cullen and Kersten 1992; Sakaguchi et al. 1992; Grassin and Fauquembergue 1996).

The pectinolytic system is composed of several enzymes distributed in two main groups: pectinesterases, which catalyze the de-esterification of methoxyl groups of pectin and depolymerases that cleave α-1,4 glycosidic linkages by hydrolysis or trans-elimination reactions (Sakai et al. 1993). Polygalacturonases (PGs) are depolymerases that specifically hydrolyze the glycosidic bonds of polygalacturonic acid.

Genes encoding PG in the *Penicillium* species have been cloned and characterized due to the use of this enzyme in the biological treatment of waste-water containing pectin and its role in infection and plant disease (Ishida et al. 1997; Wagner et al. 2000). An interesting feature of fungal PG is the presence of some conserved residues that have recently been proven to be necessary for enzyme activity (van Santen et al. 1999).

Gene cloning has opened new ways to obtain fungal strains with increased expression levels of proteins that are of industrial importance. The introduction of multiple copies of the gene of interest and gene-fusion strategies have been used to increase production of proteins such as glucosamylase, endoglucanase, xylanase, bovine prochymosin, and human interleukin-6 (Hata et al. 1991; Saarelainen et al. 1993; Sánchez-Torres et al. 1994; Tsuchiya et al. 1994; Gouka et al. 1996).

Our group has been studying the pectinolytic system of *Penicillium griseoroseum* as a model system for gene organization and regulation in filamentous fungi. Our aim is to obtain overproducing strains that can be used in the textile industry to degum natural fibers. Two genes encoding endoPGs, *pgg1* and *pgg2*, have been cloned and are differentially expressed in response to the growth medium (Ribon et al. 1999; Ribon et al. 2002). Studies are being done to investigate the molecular mechanisms that control their expression. Here we show the organization of both genes in *P. griseoroseum* and compare it to the organization of homologous DNA sequences in the genome of other *Penicillium* species. Some features found in *Aspergillus* and *Neurospora* genes are also seen in *Penicillium* as revealed by the comparison of the structural unit of the genes encoding PG cloned until now.

Materials and Methods

Fungal strains and inoculum production

*Penicillium charlesii* (CCT 4752), *P. chrysogenum*, *P. citrinum* (CCT 3281), *P. expansum*, *P. griseoroseum*...
(CCT 6241), P. italicum, P. janthinellum (CCT 3162), and P. purpurogenum (CCT 2008) were employed in this study. The stock cultures were maintained in glycerol at 4 °C. Petri dishes containing minimal medium (Pontecorvo et al. 1953) covered with cellophane membrane were inoculated with 10⁵ conidia and incubated for two days at 28 °C. The mycelia were removed from the cellophane for DNA extraction.

DNA isolation and Southern blot

Total DNA from the Penicillium species was extracted as described by Specht et al. (1982), and cleaved with EcoRI, SacI, and SalI. The reactions were analyzed on 0.8% agarose gel, and then transferred to Duralon membranes (Stratagene) according to Sambrook et al. (1989). The probe consisted of a DNA fragment amplified from total DNA from P. griseoroseum, purified from the gel and radio labelled with the Random Prime-II Labeling Kit (Stratagene). This fragment contains a 420-bp homologous domain observed in endoPG genes from several filamentous fungi and bacteria, and was generated as described by Cary et al. (1995). Hybridization was carried out overnight at 60 °C in standard hybridization buffer (Sambrook et al. 1989), washed twice with 2 X SSC, 0.1% SDS, for 20 min, and once with 1 X SSC, 0.1% SDS for 10 min. Autoradiographs were made by five-day exposure of XOMAT K film (Kodak), with an intensifying screen.

Nucleotide sequence accession number

GenBank and EMBL databases were scanned for PG-encoding genes from Penicillium species. Seven genes were retrieved and employed for the comparisons: AB015286 (P. digitatum), AF047713 (P. expansum), AF085238 and AF195790 (P. griseoroseum), D79980 (P. janthinellum), AJ243521 and AJ243522 (P. olsonii).

Results and Discussion

Genomic organization of PG-encoding genes in Penicillium griseoroseum

All endoPG genes described until now for filamentous fungi have a homologous domain of approximately 400 bp assumed to contain residues that are relevant for PG activity. A 420 bp-DNA fragment containing the corresponding domain of P. griseoroseum was used as probe in Southern blot analysis of the genomic DNA to investigate the existence of an endoPG gene family in the genome of this fungus. Few hybridizing fragments were seen even when low stringency conditions were employed during the hybridization step (Figure 1). Re-probing of the membrane with the radiolabelled pgg1 and pgg2 genes was conducted at 65 °C. The bands detected corresponded to those seen in Figure 1 confirming the presence of only two endoPG genes in the fungus genome (data not shown). The existence of exoPG genes is possible although not detected in our study due to the low similarity with the endoPG genes.

As seen in Figure 1 we can conclude that the pgg1 and pgg2 genes exist as single copies in the P. griseoroseum genome since there are no internal restriction sites for the enzymes EcoRI, SacI, and XbaI. Single copies have been demonstrated for other PG genes including the clpg1 and clpg2 genes of Colletotrichum lindemuthianum, pg1 and pg2 of P. olsonii, and bcpg1 of Botrytis cinerea (Centis et al. 1996; Centis et al. 1997; ten Have et al. 1998; Wagner et al. 2000). It also seems that pgg1 and pgg2 are present in the same copy number and that the differential levels of expression reported for these genes by Ribon et al. (2002) are probably due to gene position in the genome. The possibility that they are controlled by different cis-acting factors is now under investigation.

Occurrence of corresponding PG genes in other Penicillium species

The presence of DNA sequences homologous to endoPG genes was investigated throughout the Penicillium species under low stringency hybridization (Figure 2). All species tested have at least one endoPG gene. However, most of them showed two strongly hybridizing bands, although with some length polymorphism, that may correspond to homologous genes. As opposed to Penicillium, gene families consisting of at least five PG members have been detected in Aspergillus niger, B. cinerea, and
Sclerotinia sclerotiorum (Fraissinet-Tachet et al. 1995; Parenicová et al. 1998, Wubben et al. 1999). Wagner et al. (2000) characterized two PG genes in P. olsonii that showed significant homology to endoPG of other filamentous fungi and suggested the presence of at least one more gene in the fungus genome.

The same hybridization pattern seen in P. griseoroseum, P. expansum, P. italicum, and P. purpurogenum suggests that their PG genes have a similar global organization in the genome segment in which they are found. The same pattern observed for P. charlesii and P. citrinum could also suggest that these fungi are closely related although more studies are needed. Molecular characterization of ten Penicillium species by random amplified polymorphic DNA (RAPD) and restriction fragment length polymorphism (RFLP) of the internal transcribed spacer region of rDNA done in our laboratory revealed no difference between P. griseoroseum, P. expansum, and P. purpurogenum indicating once more a close phylogenetic relationship, and morphological and physiological characteristics confirm that they are different species (Pereira et al. 2002).

Comparison of the structural units of PG genes from Penicillium species

Although Penicillium can also be considered an important fungal model, no compilation of gene organization data from these organisms has been performed until now that could reveal important sequences for gene expression as already reported for Neurospora and Aspergillus. Here we present a comparison of the seven PG genes from P. griseoroseum, P. expansum, P. digitatum, P. janthinellum, and P. olsonii retrieved from the GenBank and EMBL databases. Their nucleotide and amino acid sequences were compared for general features such as number, length, splice sites, and position of introns, consensus sequences flanking the translation initiation codon, and codon usage.

All polypeptides started invariantly with AUG. It was not possible to define a consensus sequence surrounding the initiation codon due to the small number of PG genes available from Penicillium. However, there is a preference for purines (Pu) before ATG (Pu57Pu100Pu71ATG), with an A always located at -3. Ballance (1986) described the sequence TCACAATGGC as the consensus for the AUG environment while CAMMATGGCT (M = C, A) was defined on comparison of 88 genes from Neurospora (Edelmann and Staben 1994). The three stop codons were used at least once.

PG-encoding genes from the Penicillium species compared in this study have intervening sequences conserved in the same positions, although there was no conservation of their nucleotide sequences. The highest identity (42.1%) was found between the second introns of P. griseoroseum and P. digitatum. The average length of the introns was 61 bp and the largest one (88 bp) was observed in the PG gene from P. janthinellum. The pepg1 gene from P. expansum and the pg2 from P. olsonii showed an additional intron conserved at the same position, while the only intron detected in the pg1 gene of P. olsonii was placed with the second intron of the other genes. Conservation of the intron position, but not sequence, is usually seen be-
tween corresponding genes of filamentous fungi suggesting a common ancestor (Gurr et al. 1987; Unkles 1992). The gpgI, gpgII, gpgC, and gpgE genes from A. niger are punctuated by two, one, three, and three introns of different sequences although some of them share the same position (Benen et al. 1996).

The consensus sequence GTANRN, with preference for T in the sixth position, was found at the 5' splice site of the introns of the PG genes and resembles the sequence GTRNGY (Y = C, T) described by Unkles (1992). All 3' splice sites were represented by PyAG. The internal consensus sequence RCTRAC was described by Orbach et al. (1986) as necessary for the correct intron splicing. In this comparison, CTRACT/G was the nearest sequence that matched the one described.

Gurr et al. (1987) reported that filamentous fungi genes tend to exhibit codon preferences. The codon usage in the genes encoding PG in the Penicillium species was biased, and there was a high preference for C at the third position, followed by T, G, and A (Table 1). The codon usage seen in this study was similar to that described by Bussink et al. (1991) when the A. niger pgal and pgall genes were compared. Also, codon preference was the same as observed from direct comparison between the glaA genes from A. niger, A. oryzae, and A. shiroomi (Unkles 1992). It is believed that fungal genes with a high expression level show more bias in their choice of codons (Gurr et al. 1987). However, this does not seem to be a rule for the PG genes from P. griseoroseum since their codon usage was nearly the same (47 for pgg1 and 44 for pgg2) and a higher expression level was seen for pgg2 (Ribon et al. 2002).

Pectinases are enzymes widely used in the food industry. Genes encoding PG and their corresponding enzymes are well studied in Aspergillus, traditionally used as the model genus in filamentous fungi. However, Penicillium is an alternative that can be used to investigate aspects of gene organization and regulation due to the presence of multigenic families, genes differentially expressed and transposon-like elements, as well as its importance and potential application of the proteins and secondary metabolites in the food and pharmaceutical industry. The interesting features revealed by analysis of the nucleotide sequences can be employed in phylogenetic studies and construction of expression vectors.

Acknowledgments

This research was funded by the Fundação de Amparo a Pesquisa do Estado de Minas Gerais, grant no. CBS 1076 and the Financiadora de Estudos e Projetos/Banco Interamericano de Desenvolvimento, grant no. 64.93.0020.00.

The authors acknowledge support from a fellowship from the Conselho Nacional de Desenvolvimento Científico e Tecnológico.

References


