Available online at www.sciencedirect.com
ScienceDirect
Fungal Genetics and Biology 43 (2006) 707–714
FUNGAL GENETICS AND BIOLOGY
www.elsevier.com/locate/yfgbi

Gene duplication event in family 12 glycosyl hydrolase from *Phytophthora* spp.

Stefano Costanzo [a,c], M.D. Ospina-Giraldo [b], K.L. Deahl [c], C.J. Baker [d], Richard W. Jones [c,*]

[a] *Department of Natural Resource Sciences and Landscape Architecture, University of Maryland, College Park, MD 20742, USA*
[b] *Department of Biology, Lafayette College, Easton, PA 18042, USA*
[c] *Vegetable Laboratory, USDA-ARS, 10300 Baltimore Ave., Beltsville, MD 20705, USA*
[d] *Molecular Plant Pathology Laboratory, USDA-ARS, 10300 Baltimore Ave., Beltsville, MD 20705, USA*

Received 24 December 2005; accepted 27 April 2006
Available online 19 June 2006

Abstract

A total of 18 paralogs of xyloglucan-specific endoglucanases (EGLs) from the glycosyl hydrolase family 12 were identified and characterized in *Phytophthora sojae* and *Phytophthora ramorum*. These genes encode predicted extracellular enzymes, with sizes ranging from 189 to 435 amino acid residues, that would be capable of hydrolyzing the xyloglucan component of the host cell wall. In two cases, four and six functional copies of these genes were found in tight succession within a region of 5 and 18 kb, respectively. The overall gene copy number and relative organization appeared well conserved between *P. sojae* and *P. ramorum*, with apparent synteny in this region of their respective genomes. Phylogenetic analyses of *Phytophthora* endoglucanases of family 12 and other known members of EGL 12, revealed a close relatedness with a fairly conserved gene sub-family containing, among others, sequences from the fungi *Emericella desertorum* and *Aspergillus aculeatus*. This is the first report of family 12 EGLs present in plant pathogenic eukaryotes.
Published by Elsevier Inc.

Keywords: Xyloglucan hydrolase; Endoglucanase; *Phytophthora infestans*; *Phytophthora sojae*; *Phytophthora ramorum*

1. Introduction

Plant pathogens deploy an array of host cell wall modifying enzymes during pathogenesis (Gotesson et al., 2002). Host plants attempt to counter many of these enzymatic activities with proteinaceous inhibitors. The most thoroughly studied inhibitors have been the polygalacturonase inhibitor proteins (Cook et al., 1999; Desiderio et al., 1997; Leckie et al., 1999). For this family of proteins a specific pathogen endopolygalacturonase counterpart has been clearly identified (Have et al., 1998). Recently a new class of inhibitor has been identified, the xyloglucan-specific endoglucanase inhibitor proteins (XEGIP) (Naqvi et al., 2005; Qin et al., 2003). This inhibitor was first identified in tomato, and we have cloned and characterized XEGIP in potato. Our EST surveys indicate XEGIP is widespread among various plant genera. The counterpart to XEGIP is found only in family 12 glycoside hydrolases, however, there are no published reports on the occurrence of family 12 endoglucanases in plant pathogens. Several glycoside hydrolase family 12 endoglucanases have been identified and characterized from various bacterial and fungal saprophytes, and can be found in the CAZy Carbohydrate-Active Enzymes database (http://afmb.cnrs-mrs.fr/CAZY/). In an effort to define what role XEGIP could play in host defense from pathogens in the genus *Phytophthora*, we initiated a study to determine if family 12 endoglucanase-encoding genes occur in species of *Phytophthora*. We scanned and translated public genomic and EST databases using three conserved amino acid motifs found in family 12 endoglucanases (Goedegebuur et al., 2002). In this study, we report of a total of 18 family 12

* Corresponding author. Fax: +1 301 504 5555.
E-mail address: jonesr@ba.ars.usda.gov (R.W. Jones).

1087-1845/$ - see front matter Published by Elsevier Inc.
doi:10.1016/j.fgb.2006.04.006

endoglucanase genes identified in *Phytophthora sojae* and *Phytophthora ramorum* genomes, four of which carry an additional domain resembling a bacterial cell wall surface anchor domain. Preliminary results also show the presence of at least three copies of family 12 EGL in *Phytophthora infestans*.

2. Materials and methods

2.1. Culturing of Phytophthora spp.

Three *Phytophthora* species were used in this study: *P. infestans*, *P. sojae* and *P. ramorum*. A *P. infestans* isolate was obtained from the collection maintained at the Vegetable Laboratory, Beltsville Agricultural Research Center. The *P. sojae* isolates race 1, 4, 7 and 25 were kindly provided by Dr. S. Li at the National Soybean Pathogen Collection Center, Department of Crop Sciences, University of Illinois. All isolates were routinely cultured on rye agar medium supplemented with 2% sucrose (Caten and Jinks, 1968). Genomic DNA and total RNA from *P. ramorum* was kindly provided by Dr. M. Palm and Dr. M.C. Aime in the Systematic Mycology and Botany Laboratory at the Beltsville Agricultural Research Center.

2.2. Nucleic acid manipulation

Genomic DNA and total RNA were extracted from *Phytophthora* mycelium grown in pea broth (filtrate from 120 g of autoclaved frozen peas, supplemented with 2 g of calcium carbonate and 0.05 g of β-sitosterol per liter of medium) for approximately two weeks at 22 °C in the dark before harvesting. Total RNA was extracted using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA) following the protocol supplied by the manufacturer. Genomic DNA was extracted using the GenElute Plant Genomic DNA Miniprep Kit (Sigma, St. Louis, MO), following the instructions provided by the supplier.

2.3. RT-PCR

For gene expression analysis by RT-PCR, 5 μg of DNase-treated (TURBO DNA-free, Ambion, Austin, TX) total RNA was reverse-transcribed to cDNA (SuperScript III One-Step RT-PCR Kit, Invitrogen) with gene-specific primers (Table 2). A control reaction (minus RTase) was included to exclude spurious amplifications due to potential presence of contaminating DNA. An additional internal control for in-planta RT-PCR experiments was obtained by designing intron-spanning primers, based on the gene model of the cellulose synthase catalytic subunit located 40 kb upstream of the major EGL 12 gene cluster in the *P. sojae* genome. Forward and reverse primer sequences for the internal control are 5′-CTCGGGTTCCTCTACTAC-3′ and 5′-TCGTTGTCCAGAATGATG-3′, respectively. The expected amplicon size is 702 bp including the 94 bp intron and 608 bp without it. Annealing temperature for this primer set is 58 °C. Primers for PCR amplification of the complete open reading frame (ORF) of each EGL paralog were designed using the Primer3 web site (Rozen and Skaletsky, 2000), based on the *P. sojae* and *P. ramorum* genomic sequence data from the Joint Genome Institute. Amplification products were analyzed by 1% agarose gel electrophoresis. RT-PCR products were excised from the gel or directly purified using PureLink PCR Purification Kit (Invitrogen) following the manufacturer's protocol. The purified DNA amplicons were cloned into pCR4-TOPO vector (Invitrogen–Life Technologies). Transformed *E. coli* colonies were screened by PCR and positive clones were sequenced.

2.4. Sequence scans and analysis

Preliminary tBlastP analyses of ESTs in GenBank were performed using the conserved amino acid sequences for family 12 consensus domains, referred to as Box 1 (NNLWG), Box 2 (YELMIW) and Box 3 (GTEPF) (Fig. 2). Overlapping ESTs were identified resulting in a family 12 endoglucanase gene of expected size. This gene was subsequently used to identify *P. sojae* and *P. ramorum* sequences by BlastN analysis (*E* value $1\,e^{-5}$), using the version 1.0 of their complete genome data produced by the US Department of Energy Joint Genome Institute (http://www.jgi.doe.gov/). Phylogenetic tree construction and molecular evolutionary analyses were conducted using *MEGA* version 3.0 (Kumar et al., 2004). A phylogenetic tree was derived based on the Neighbor-Joining algorithm, including protein sequences of 49 representatives of the glycoside hydrolase family 12 among those present in the CAZy Carbohydrate-Active Enzymes database (Coutinho and Henrissat, 1999). Bootstrap test with 1000 replications were conducted to examine the reliability of the interior branches (Felsenstein, 1985). Selective pressure on individual sites of codon alignments was determined by the molecular evolution analysis program Datamonkey (http://www.datamonkey.org). The random effects likelihood (REL) analysis (Pond and Frost, 2005) was implemented to test for purifying or diversifying selection. Amino acid sites under selective pressure were identified based on a Bayes factor of 95. Alignments were formatted using *BOXSHADE* version 3.21, available online from the Swiss Institute of Bioinformatics. Motif and signal peptide predictions were carried out utilizing software developed by the Center for Biological Sequence Analysis at the Technical University of Denmark, DTU (http://www.cbs.dtu.dk/). Sequences of all the endoglucanase genes from *P. sojae* and *P. ramorum* were deposited in GenBank Accession Nos. AY645943, DQ206888–DQ206903 and DQ286231.

2.5. Infection assay

Soybean seeds (Glycine Max (L.) Merr.) of cv 'Williams' were germinated in moistened paper towels. Inoculation of