Available online at www.sciencedirect.com
SCIENCE DIRECT®
Fungal Genetics and Biology 40 (2003) 197–206
FUNGAL GENETICS AND BIOLOGY
www.elsevier.com/locate/yfgbi

Characterization of the glucose-6-phosphate isomerase gene in *Phytophthora infestans* reveals the presence of multiple alleles

M.D. Ospina-Giraldo and R.W. Jones*

USDA/ARS/Vegetable Laboratory, BARC-WEST Building 010A/Room 309, 10300 Baltimore Avenue, Beltsville, MD 20705, USA

Received 11 February 2003; accepted 17 July 2003

Abstract

Glucose-6-phosphate isomerase (GPI) plays a key role in both glycolysis and gluconeogenesis. Isoforms of GPI are common, and therefore, its isozyme pattern is widely used to characterize isolates of *Phytophthora infestans*. Despite the importance of GPI in *P. infestans* studies, the gene encoding this enzyme has not yet been characterized. Furthermore, it has been suggested that *P. infestans* contains multiple copies of the gene but this hypothesis remains to be demonstrated. We have cloned and characterized *GPI* in various isolates of *P. infestans* as well as in several species of the genus *Phytophthora*. The gene contains 1671 bp and encodes a protein with a predicted molecular weight of 60.8 kDa. Multiple different alleles were identified and Southern analysis indicated certain *P. infestans* isolates carry several copies of the gene. Phylogenetic analysis revealed that *P. infestans* GPI is most closely related to sequences from *Toxoplasma gondii*, *Arabidopsis thaliana*, and *Clarkia lewisii*.

Keywords: Gene expression; Genome analysis; Glucose-6-phosphate isomerase; GPI; Oomycetes; Pathogenicity; PGI; Phosphoglucoisomerase; *Phytophthora infestans*

1. Introduction

Potato late blight disease, caused by the oomycete *Phytophthora infestans*, represents a major agricultural challenge not only in the United States but also in those parts of the world where potatoes constitute a primary source of food. The appearance of new, more virulent races of *P. infestans* during the last decade (Fry and Goodwin, 1997) poses an increasingly significant threat to the cultivation of this staple crop. Recently, potato late blight has wreaked havoc in countries such as Russia, where it has caused losses that in some cases exceed 70% of the crop (Garelik, 2002). The simultaneous occurrence of multiple genotypes that respond in different ways to the control methods currently available makes accurate identification an indispensable requirement (Goodwin et al., 1995b). Thus far, isozyme characterization has been one of the most frequently used methods to identify isolates of *P. infestans*. Polymorphic loci encoding glucose-6-phosphate isomerase (GPI), peptidase (PEP), malic enzyme (ME), and xanthine dehydrogenase (XDH) have been successfully utilized to unequivocally differentiate multiple individuals within a sexual population (Tooley et al., 1985) but it is in the GPI and PEP loci where polymorphisms are most commonly observed.

Currently, the GPI isozyme pattern is the most widely used isozyme approach to characterize isolates of *P. infestans*, in which the GPI isozymes occur in multiple combinations, as shown by starch gel and cellulose-acetate electrophoresis analyses (Goodwin et al., 1995a; Tooley et al., 1985). Previous studies have revealed that at least 6 different genotypes can be found in sexual populations of *P. infestans* (Tooley et al., 1985), although many more are possible (Grunwald et al., 2001). These genotypes are known as 83/100, 86/86, 86/100, 100/100, 100/122, and 122/122. (Each allele is named according to the electrophoretic mobility of the isozyme it encodes, in relation to the most common one, which is known as 100). Until the late-1980s, the population of *P. infestans* in North America was mostly represented by single clonal lineages known as US-6,

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

1087-1845/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S1087-1845(03)00107-5

which carried the 100/100 genotype, and US-1, which carried the 86/100 genotype (Goodwin et al., 1994; Tooley et al., 1985). In the early 1990's two additional lineages, US-7 and US-8 (GPI genotypes 100/111 and 100/111/122, respectively) were also found in the United States and Canada (Goodwin et al., 1995b). In addition, GPI genotype 100/122, found in US-14 and US-17 lineages, has become increasingly predominant while genotype 100/100/111 (characteristic of US-11), frequently observed between 1995 and 1997, significantly decreased by 1998 (Daayf and Platt, 2000). Although most of these genotypic changes have been attributed to migration rather than mutation (Goodwin et al., 1994), the widespread occurrence of the newest genotypes might suggest a selective advantage for such genotypes and perhaps, a role for sexual recombination.

Glucose-6-phosphate isomerase, also known as phosphoglucose isomerase or PGI, catalyzes the interconversion of D-glucose-6-phosphate and D-fructose-6-phosphate (Achari et al., 1981), and therefore, constitutes an integral part of both glycolysis and gluconeogenesis. Prokaryotic and eukaryotic GPIs are dimeric enzymes, with two catalytic sites composed of parts of each subunit. Both subunits are essentially identical and each monomer has a molecular mass of approximately 60–62 kDa (Achari et al., 1981; Froman et al., 1989; Green et al., 1988; Kaslow and Hill, 1990; Thomas et al., 1993).

Despite the crucial role of GPI in *P. infestans* characterization studies, the gene encoding this enzyme has not been characterized yet. Furthermore, it has been suggested that *P. infestans* contains multiple copies of the *GPI* coding sequence (Goodwin et al., 1995a) but this hypothesis remains to be demonstrated. In an attempt to elucidate these questions, we have cloned and characterized the *GPI* gene in various isolates of *P. infestans* as well as in several other species of the genus *Phytophthora*.

2. Materials and methods

2.1. Isolates

All samples used in this study are part of the *Phytophthora* spp. collection maintained at the Vegetable Laboratory, Beltsville Agricultural Research Center, and were kindly provided by K. Deahl (Table 1). Fresh cultures were 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) at 22 °C in the dark.

2.2. Nucleic acid manipulations

Total RNA from *Phytophthora* mycelium grown in pea broth 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. Cloning of the *GPI* gene was done by a two-step process. In the first

Table 1
List of species, isolate code, geographic origin, GPI profile, clone ID, and GenBank accession number of sequences analyzed in this study

Species	Isolate code	Geographic origin	GPI profile	Clone ID	Accession
P. infestans	EC3215	Ecuador	86/100	MO41	AY220245
P. infestans	EC3215	Ecuador	86/100	MO42	AY220246
P. infestans	EC3215	Ecuador	86/100	MO43	AY220247
P. infestans	EC3215	Ecuador	86/100	MO44	AY220248
P. infestans	EC3215	Ecuador	86/100	MO45	AY220249
P. infestans	EC3215	Ecuador	86/100	MO46	AY220250
P. infestans	FL-01-3	United States	100/111/122	MO12	AY098641
P. infestans	FL-01-3	United States	100/111/122	MO35	AY220240
P. infestans	FL-01-3	United States	100/111/122	MO36	AY220241
P. infestans	Indo-2	Indonesia	86/86	MO37	AY220242
P. infestans	MX550	Mexico	86/86	MO20	AY220260
P. infestans	Phil-1	Philippines	86/100	MO15	AY220258
P. infestans	US97-001	United States	100/122	MO40	AY220244
P. infestans	WL98-139	Wales	100/100	MO47	AY220251
P. infestans	WL98-139	Wales	100/100	MO48	AY220252
P. infestans	WL99-1-1	Wales	100/100	MO17	AY220259
P. cactorum	PC1	United States	ND	MO49	AY220253
P. cactorum	PC1	United States	ND	MO51	AY220254
P. erythroseptica	PE99 23-3 MR	United States	86/86	MO53	AY220256
P. erythroseptica	PE99 23-3 MR	United States	86/86	MO55	AY220257
P. erythroseptica	PE99 18-5 MS	United States	86/86	MO52	AY220255
P. mirabilis	PM64072	Mexico	ND	MO38	AY220243

MR, metalaxyl-resistant; MS, metalaxyl-sensitive; and ND, not determined.

stage, approximately 1 kb of *GPI*'s 5′-end transcribed sequence was amplified by reverse transcription polymerase chain reaction (RT-PCR) using the GeneRacer Kit (Invitrogen–Life Technologies, Carlsbad, CA) according to the manufacturer's directions (see Table 2 for sequences and annealing sites of all primers used in this study). Reverse transcription was done with 5 μg of total RNA and primer GPIR1, using Superscript II reverse transcriptase. Subsequent PCR was performed in a final volume of 25 μl, using 1 μl of cDNA and primers GPIR2 and GeneRacer 5′ or GeneRacer 5′ Nested Primer. PCR was carried out in a GeneAmp PCR System 9600 (Perkin-Elmer, Norwalk, CT), using *Taq* DNA polymerase (Promega, Madison, WI), according to the following parameters: 94 °C/30 s, 55 °C/30 s, 72 °C/30 s (one cycle); 92 °C/15 s, 55 °C/15 s, 72 °C/30 s (34 cycles), and 72 °C/5 min (one cycle). In the second stage, the remainder of the gene was amplified using a similar RT-PCR approach. cDNA was synthesized using oligo(dT) primer and PCR was performed with primers FPG and GeneRacer 3′ or GeneRacer 3′ Nested Primer. PCR products were cloned into pCR4-TOPO vector (Invitrogen–Life Technologies). Amplification of the regions upstream and downstream from the 5′ and 3′ ends of the gene, respectively, were done by TAIL-PCR (Liu and Whittier, 1995), using specific primer GPIR6 and arbitrary primer OP-T03 for the upstream region, and specific primer GPIF14 and arbitrary primer OP-U13 for the downstream region. Both products were cloned into pCR4-TOPO vector (Invitrogen–Life Technologies).

Gene expression analysis was conducted by RT-PCR, using 5 μg of DNase-treated total RNA and primer pairs GPIF18/GPIR6 and GPIF19/GPIR6. PCR conditions were as described above. For Southern analysis, 10 μg of genomic DNA digested with *Bam*HI, *Bgl*II or *Eco*RI was separated on 0.8% agarose gel and transferred onto Hybond-N^+ hybridization membranes (Amersham–Pharmacia Biotech, UK). Probe was prepared by PCR amplification of *GPI* from genomic DNA, using primers GPIF9 and GPIR8, and fluorescently labeled with the Renaissance Random Primer Fluorescein Labeling Kit (NEN Life Science Products, Boston, MA). Hybridization was carried out O/N at 65 °C using the modified Church and Gilbert buffer (Church and Gilbert, 1984). Membranes were washed in 2× SSC, 1.0% SDS, and 0.2×, 0.1% for 15 min each at 65 °C. Detection was done by chemiluminescence following the protocol suggested in the labeling kit. Membranes were exposed to Biomax MS X-ray film (Kodak, Rochester, NY).

2.3. Sequence analysis

DNA sequencing reactions were performed using the BigDye terminator cycle sequencing kit, version 3.1 (Applied Biosystems, Foster City, CA) and analyzed using an ABI3100 Sequencer (Applied Biosystems). PCR fragments and cloned products (between four and six clones per PCR product) were sequenced in both directions, using four internal primers (in addition to the amplification primers) in each direction, to ensure sequencing reliability and reduce the effect of potential variation introduced by PCR experiments. Plasmids were also sequenced with M13 primers, and all electropherograms were visually inspected for miscalled bases. Coding-region and amino acid sequence data were analyzed with the Lasergene software package (DNASTAR, Madison, WI). Sequence alignment was done with

Table 2
Primers used for PCR amplification and sequencing of *PIGPI* and adjacent regions

Primer name	Primer sequence	Location[a]
FPG	TAGCTACCTCGGCCCC	1385
GPIF4	TGCGCAAAGTTCGTAGACCCACAC	898
GPIF6	CACGAGGTGCTGGACGCCA	1270
GPIF7	CTGTGCCACTTGTGCAGCA	1684
GPIF8	GAGTCGCAGAACCCCGTG	2169
GPIF14	TCCTCCACCGGCCAGC	2379
GPIF16	CCGCTTACAATTTACGCAATCG	335[a]
GPIF17	CTTGGTTATGTCAACAATGTGTATT	686
GPIF18	AATTTGCGCTCAAGCTCGC	549
GPIF19	AATTTGCGCAAAGTT	890
GPIR1	GTCCTCCATGTCCACCTGCTG	2018
GPIR2	ATGGCATGAGCACCAGCCAGG	1849
GPIR4	CACGCACCTGCTTAGCCAGCA	2505
GPIR6	GATGGCGTCCAGCACCTCGT	1295
GPIR7	GAGCCTTCCAAGCGGGTGTG	963
GPIR8	ACTATATGAGCACATACTCGAC	2669
GPIR9	CCTGCAGCTTCTTCTTGAGC	1155
OP-T03	TCCACTCCTG	1
OP-U13	GGCTGGTTCC	3794

[a] Annealing sites (5′ end) are with respect to MO36 sequence except GPIF16 and GPIF18, whose location is in reference to MO35.

Clustal W (Thompson et al., 1994). No manual adjustments were required. Nucleotide distances were estimated by the Kimura two-parameter model (Kimura, 1980) and phylogenetic inference was performed by the distance-based Neighbor-Joining (NJ) algorithm (Saitou and Nei, 1987). Bootstrap tests with 1000 replications (Felsenstein, 1985) were conducted to examine the reliability of the interior branches and the validity of the trees obtained. All phylogenetic and molecular evolutionary analyses were conducted using MEGA (Arizona State University, Tempe, AZ). Sequence similarity was determined using BLASTN or BLASTX (Altschul et al., 1997) and motifs were identified using PSORT, a computer program for the prediction of protein localization within cells (Nakai and Kanehisa, 1992), and NetStart, a program for predicting translation start sites (Pedersen and Nielsen, 1997). Potential phosphorylation and glycosylation sites were predicted using NetPhos (Blom et al., 1999) and NetOGlyc (Hansen et al., 1998).

2.4. Protein expression analysis

Expression of the *P. infestans* gene in a bacterial system was achieved by cloning the *GPI* coding sequence into pCRT7/NT/-TOPO TA Expression vector (Invitrogen–Life Technologies) following the manufacturer's directions. Detection of the recombinant protein was done by Western blotting using horseradish peroxidase (HRP)-conjugated antibodies directed against the HisG epitope. Native or recombinant GPI activity was detected by cellulose–acetate electrophoresis (Goodwin et al., 1995a) using 10 μl of sonicated bacterial culture.

3. Results

3.1. Cloning of the GPI gene and adjacent regions

The design of primers GPIR1 and GPIR2 was based on a conserved region found in a *P. infestans* EST reported by the Phytophthora Genome Consortium of the National Center for Genomic Resources (http://www.ncgr.org). In the first cloning step, a fragment of 972 bp, containing the 5′ end of the gene, was obtained. This product was cloned into pCR4-TOPO vector, to give clone MO01. MO01 sequence was used to design FPG, a specific primer that, when used for PCR in conjunction with oligo(dT) and an oligo(dT)-primed cDNA, generated the 3′ end of the gene (937 bp, including 103 bp of overlapping sequence). The two-step cloning process of the *GPI* transcript produced a cDNA sequence of 1806 bp that included two untranslated regions of 55 and 80 bp at the 5′ and 3′ ends, respectively. The length and sequence of the cDNA's coding region were confirmed after synthesizing a new full-length cDNA by RT-PCR using specific primers GPIF4 and GPIR8, which are complementary to sections of the 5′ and 3′ untranslated regions, respectively.

Primers GPIF9 (whose 5′ end is complementary to the translation initiation site) and GPIR8 were used to amplify by PCR the *GPI* gene, using as template genomic DNA from *P. infestans* isolates indicated in Table 1. Sequence analyses of the PCR products and the respective clones revealed that the *GPI* gene, termed *PIGPI*, has a single open reading frame (ORF) of 1671 bp in length and contains no introns (Fig. 1). The ORF encodes a protein (PIGPI) that consists of 556 amino acids and has a predicted molecular weight of 60.8 kDa and a calculated isoelectric point of 6.2.

The TAIL-PCR experiments produced additional fragments of 1295 bp with primers GPIR6/OP-T03 and 1430 bp with primers GPIF14/OP-U13, corresponding to regions upstream and downstream from the 5′ and 3′ ends, respectively. Sequence assembly produced a single contig of 3.8 kb (Fig. 1). Sequence analysis of the fragment amplified with primers GPIF14/OP-U13 revealed the presence of a 656-bp putative ORF starting 534 bp downstream from the *PIGPI* stop codon (Fig. 1). A comparison of the amino acid sequence deduced from this ORF with NCBI's Conserved Domain database produced a significant alignment (*E* value: 4e − 06) with the conserved domain of a group of methylases involved in coenzyme metabolism. The fragment amplified with primers GPIR6/OP-T03 had no significant similarities with sequences reported in the databases.

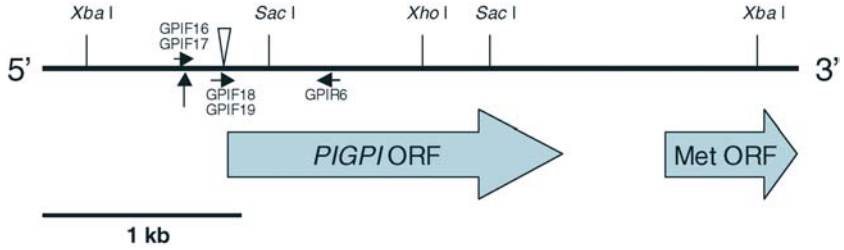

Fig. 1. Map of the 3.8-kb contig containing the GPI gene (allele I) from *P. infestans*. The arrow and the open triangle indicate the locations of the 26 bp and the 10-bp gaps, respectively. Also indicated are the primers used to specifically amplify either allele I (GPIF17 or GPIF19 and GPIR6) or allele II (GPIF16 or GPIF18 and GPIR6). Met, Methylase-like.

3.2. Sequence analysis

PIGPI's transcriptional start site is located within the predicted motif 5′-GCTCATTYYNCAWTTT-3′ (McLeod et al., 2003; Pieterse et al., 1994). Two CT-elements are present in the *PIGPI* sequence. The first motif (CTCATTC) spans the transcription initiation site while the second one (CTCTTCTT) is situated four nucleotides upstream the translation start point. A predictive analysis of protein sorting signals and localization sites using PSORT (Nakai and Kanehisa, 1992) showed that PIGPI appears to contain neither nuclear localization signals nor N-terminal signal peptide and is more likely (73.9%) to be localized in the cytoplasm. PIGPI is predicted to have multiple potential phosphorylation sites. No significant prediction was made for potential glycosylation modifications.

Based on the sequence obtained from the DNA fragment amplified by TAIL-PCR with primers GPIR6 and OP-T03, we designed internal primers that allowed us to amplify, and subsequently sequence, a smaller fragment containing the promoter and part of the 5′ end of the gene in the FL-01-3 isolate (GPI profile 100/111/122). Sequence analysis of a fragment amplified with primers GPIF11 and GPIR9 revealed the existence of two different alleles, one of which (type I, clone MO36) had two gaps of 10 and 26 bp located at nucleotide positions −30 and −229, respectively, in relation to the translation initiation codon (Fig. 1). Neither of these gaps was found in the second allele (type II, clones MO12 and MO35. These two clones have differences in the coding sequence but their promoter regions are almost identical). Sequence analysis of PCR amplicons obtained from Welsh and Indonesian isolates (genotypes 100/100 and 86/86, respectively) suggested that their *PIGPI* alleles lack the gaps found in the Florida isolate. Based on this sequence distinction, we designed four PCR primers (Table 2) that, when used in combination with any GPI reverse primer specifically amplify either type I (GPIF17 and GPIF19) or type II allele (GPIF16 and GPIF18), and confirmed the simultaneous presence of both alleles in isolates FL-01-3 and US97-001 (Fig. 2A). Indonesian and Welsh isolates analyzed with these primers appear to carry type II allele, only (Fig. 2A).

To analyze the transcriptional activity of alleles I and II in isolates FL-01-3 and US97-001, we performed an RT-PCR experiment with primer pairs GPIF18/GPIR6 and GPIF19/GPIR6. Results indicate that both alleles are transcriptionally active but the expression level of type II allele appears to be higher than that of type I allele (Fig. 2B).

A Southern analysis of restricted genomic DNA from three isolates with different GPI profile was conducted, using as probe a 1.7-kb fragment amplified by PCR from isolate FL-01-3, with primers GPIF9/GPIR8. Results suggested that *PIGPI* is present once in isolates with GPI profile 86/86, twice in isolates with profiles 100/100, and at least three times in isolates with profiles 100/111/122 (Fig. 3).

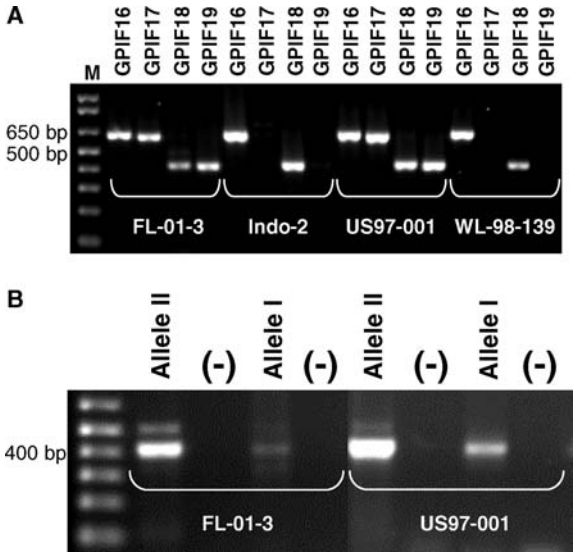

Fig. 2. Analysis of *PIGPI* from *P. infestans* isolates with different isozyme profiles. (A) PCR amplification of *PIGPI* alleles using genomic DNA. Each isolate's DNA was analyzed with either one of the forward primers indicated and the reverse primer GPIR6. (B) RT-PCR using total RNA from isolates FL-01-3 and US97-001 and primer pairs GPIF19/GPIR6 for allele I or GPIF18/GPIR6 for allele II. Dashes indicate negative controls.

3.3. Expression of PIGPI in a bacterial system

A PCR fragment obtained with primers GPIF9/GPIR8 was cloned into pCRT7/NT TOPO-TA expression vector (Invitrogen–Life Technologies). After performing a sequencing reaction to ensure that *PIGPI* had been cloned in the correct frame, expression of the recombinant fusion protein was induced as suggested by the vector supplier. A band of approximately 64 kDa (including approximately 4 kDa of the N-terminal tag), showing increased intensity, was observed in the polyacrylamide gel (Fig. 4A). Western blotting and hybridization with anti-HisG-HRP antibodies confirmed the expression of the recombinant protein (Fig. 4B). Recombinant PIGPI appeared to be fully functional as indicated by enzymatic and kinetic assays (Fig. 4C).

3.4. Phylogenetic analysis

The *GPI* gene was also cloned from *P. cactorum, P. erythroseptica* (one isolate each of metalaxyl-sensitive and resistant strains), and *P. mirabilis*. Two distinct alleles were identified in *P. cactorum* whereas three different ones were found in *P. erythroseptica*. All coding regions and deduced amino acid sequences were compared in separate analyses and, in both cases, sequence

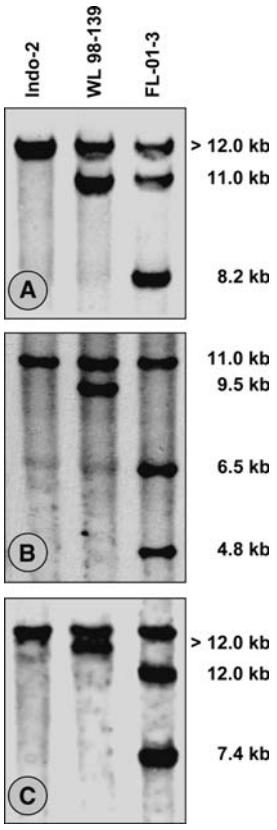

Fig. 3. Southern hybridization analysis of the *GPI* gene copy number in *P. infestans*. Restriction enzyme digests were done with *Bam*HI (A), *Bgl*II (B), and *Eco*RI (C). Numbers on the right indicate the approximate size (in kilobases) of fragments hybridizing to a GPI probe obtained by PCR with primers GPIF9 and GPIR8. The GPI profile for each isolate is described in Table 1.

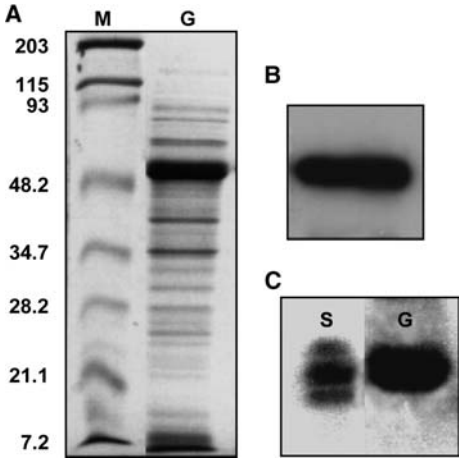

Fig. 4. Analysis of recombinant GPI in a bacterial system. (A) Polyacrylamide gel showing a 64-kDa band with increased intensity. (B) Western blotting and hybridization with anti-HisG-HRP antibodies confirmed the in vitro expression of the recombinant GPI protein. (C) Recombinant PIGPI appeared to be fully functional as indicated by an enzymatic assay performed directly on cellulose-acetate electrophoresis. M, molecular mass standards (kDa); G, recombinant GPI; and S, US-8 GPI standards.

variation was observed among all but three of the samples, with a total of 86 polymorphic nucleotide sites found in the coding sequence. Isolates Phil-1, US97-001 and one sequence (MO12, a variation of the type II allele) generated from FL-01-3 had identical nucleotide sequences. The highest percentage of nucleotide substitutions among the studied sequences was 1.97%, observed between the metalaxyl-susceptible isolate of *P. erythroseptica* and the Indo-2 isolate of *P. infestans*. Most of these changes were transitions and the average transition/transversion ratio was 4.111. At the amino acid level, a total of 47 variable sites were found and sequence comparison indicated that the maximum variation was 3.2%, observed between the same isolates of *P. erythroseptica* and *P. infestans* described above. Overall, the average amino acid variation was 1.17%.

Comparisons among deduced amino acid sequences corresponding to the *GPI* alleles amplified from *P. infestans* isolates with different GPI isozyme profiles showed that the percentage of amino acid substitutions ranged within 0.4% [between genotypes 86/100 and 100/111/122 (allele I)] and 2.3% [between genotypes 86/86 and 100/100 (Table 3)]. In addition, we compared the nucleotide sequences obtained from all *P. infestans* isolates with a *PIGPI* consensus sequence and computed the number of synonymous and non-synonymous substitutions. The average ratio of non-synonymous versus synonymous substitutions was 1.58 and the highest value was 4.0, observed in sequences MO35 (type II allele of FL-01-3) and MO47 (isolate WL98-139). Polymorphic sites at nucleotide and amino acid level occurred randomly within the analyzed sequences (although a region spanning nucleotide sites 571 through 685 contained no variation at all) and no evidence of the presence of a highly variable region, or 'hot spot', was observed. Hard copies of the sequence alignments are available upon request.

BLASTX analysis of the *PIGPI* consensus sequence indicated that it is most similar to homologs from

Table 3
Maximum percentage of amino acid substitutions between GPI protein sequences from isolates with different GPI isozyme profiles

	86/86	86/100	100/100	100/111/122 allele I
86/100	1.4			
100/100	2.3	0.9		
100/111/122 allele I	1.8	0.4	1.3	
100/111/122 allele II	2.2	0.7	1.6	1.1

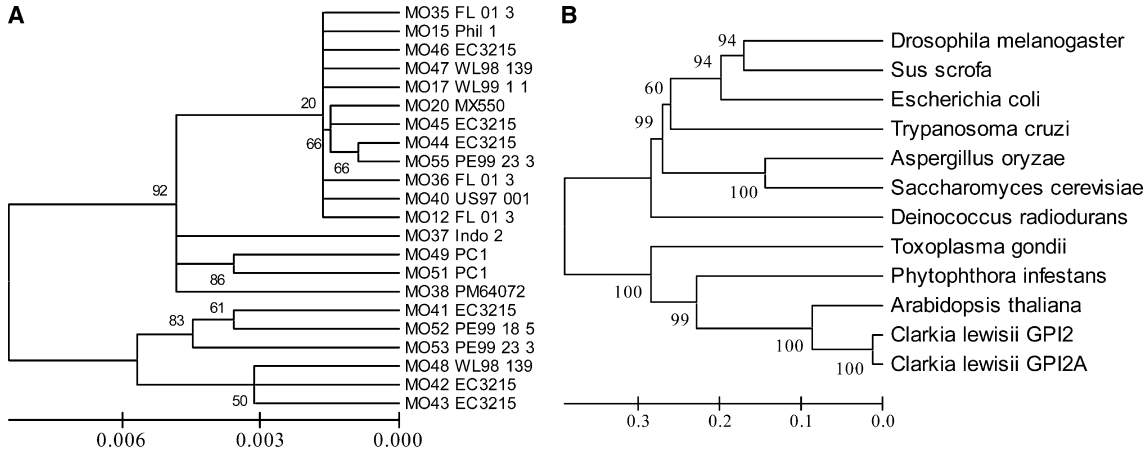

Fig. 5. (A) Phylogenetic analysis of 22 alleles of the *GPI* gene from *P. infestans* and other species of the genus *Phytophthora*. (B) Phylogenetic relationships of the *P. infestans* GPI consensus sequence with 11 homologs from members of other Kingdoms. The phylogenetic trees were obtained by the Neighbor-Joining method and the number of amino acid residues in each subset was 556 (A) and 526 (B). There were no equivocal sites and gaps were deleted from the datasets. The numbers indicate the values obtained after a bootstrap test with 1000 replications. The scale bar represents the branch length measured in the number of amino acid substitutions per site.

Arabidopsis thaliana and *Clarkia* spp., having 63 and 62%, respectively, identical amino acid residues. At the nucleotide level, the highest similarity is with *GPI* genes from the bacteria *Deinococcus radiodurans* and *Magnetospirillum magnetotacticum*, and the eukaryote *Toxoplasma gondii*.

Molecular evolutionary analyses were conducted with nucleotide (coding-region only) and amino acid sequences and both types of phylogenetic trees obtained show very similar topology. The protein-based phylogenetic tree is comprised of two groups, a distribution supported by a bootstrap value of 92%. Sixteen of the 22 samples analyzed in this study grouped together in a major cluster (Fig. 5A), with sequences from *P. cactorum*, *P. mirabilis*, and the metalaxyl-resistant isolate of *P. erythroseptica* being part of this subset. The tree includes a smaller, sister cluster constituted by three alleles of the Ecuadorian isolate, one allele of the Welsh isolate 98-139, and one allele each of the metalaxyl-resistant and metalaxyl-susceptible isolates of *P. erythroseptica*. A consensus sequence was deduced from the 16 different *P. infestans* GPI sequences analyzed in this study and compared with other GPI protein sequences from representative prokaryotic and eukaryotic sources available in the databases. Phylogenetic analyses revealed that the *P. infestans* consensus GPI shared the most recent ancestor with sequences from *Clarkia lewisii* and *A. thaliana*. Both groups, in turn, shared the most recent ancestor with the *T. gondii* GPI sequence (Fig. 5B).

4. Discussion

Frequently, genes playing central roles in particular metabolic processes remain uncharacterized, even though they can be of significant importance in the cell's evolution (Raymond et al., 2002). Despite the importance of glucose-6-phosphate isomerase in the detection and identification of *P. infestans*, the gene encoding this enzyme has not been described in the literature nor has its full sequence been reported in the public databases. Based on a short EST containing a putative *GPI* sequence, we designed primers that allowed us to generate a full-length cDNA by RT-PCR, using a common approach to amplify cDNA ends. Ultimately, we were able to clone and characterize an approximately 3.8-kb long DNA fragment that contains the *P. infestans* GPI gene and flanking regions extending for more than 1 kb outside the coding sequence.

Changes in GPI could be an important factor in pathogen variation. For example, physiological differences among GPI genotypes have been found in several populations of the willow beetle *Chrysomela aeneicollis* suggesting that variation in the GPI genotype might be correlated to temperature adaptation (Dahlhoff and Rank, 2000). Moreover, GPI has been shown to play a role in pathogenicity by *Xanthomonas campestris* (Tung and Kuo, 1999), a finding that prompted the question whether a similar role could be played in *P. infestans*. In addition, isozyme studies would suggest that the GPI profile has undergone a considerable evolutionary variation that is concurrent with the appearance of new, more virulent isolates.

Isozyme analysis relies on the fact that differences in amino acid composition produce variation in the net charge of nearly identical proteins, and consequently, in the protein migration rate on an electrical field. In fact, our computer calculation of the electrical charge of each of the deduced GPI proteins analyzed in this study show that at pH 8.5, the pH at which the cellulose-acetate electrophoresis analysis is conducted, all deduced proteins have different net charge, ranging from −7.8 to

−11.8. However, certain GPI banding patterns observed after isozyme analysis still remain to be fully explained. For example, the presence of the extra isozyme bands detected in isolates with GPI profiles 100/111/122 or 100/100/111 has been suggested to be the result of an extra chromosome in the *P. infestans* genome (Goodwin et al., 1995a), but the lack of specific DNA probes has prevented testing such a hypothesis. In this work, we provide evidence obtained from Southern hybridization, sequencing and PCR experiments, suggesting that multiple copies of *PIGPI* do exist, and are transcriptionally active, in some isolates of *P. infestans* (Figs. 2 and 3). Because the banding pattern in a Southern hybridization depends on the extent of polymorphism existing on the chromosomal region being probed, it is difficult to ascertain whether the additional *PIGPI* copies are the result of the presence of an extra chromosome or simply correspond to a diverged duplicated region. Our results are consistent with either alternative and, therefore, we cannot rule out the possibility that the banding pattern observed in the Southern analysis is the result of a restriction enzyme length polymorphism between the existing and the additional chromosomes. Conversely, *GPI* duplication has been reported in several species of the genus *Clarkia* (Thomas et al., 1993) and a similar phenomenon could have led to the appearance of multiple copies of *GPI* in the same chromosome of *P. infestans*, an organism in which gene duplication has previously been described (Pieterse et al., 1994).

The *GPI* gene sequence lacks introns, a feature common to most *P. infestans* genes characterized so far, with *PiYpt1*, a gene encoding a RAS-related protein involved in vesicle transport (Chen and Roxby, 1996), and *NiaA*, a gene encoding nitrate reductase (Pieterse et al., 1995), being notable exceptions. Other sequence features common to *P. infestans* genes, such as context of the transcription initiation point and potential promoter signals, are also present in *PIGPI*. The CT elements, which are thought to play a role in determining the frequency of transcription initiation in *P. infestans* (Chen and Roxby, 1997), lie just outside a transcribed region that, in the type I allele cloned from the Florida isolate, contains a 10-bp gap. The presence of such a gap could influence the context in which the CT-elements work. In fact, allele type I appears to be expressed at a lower rate than allele type II (Fig. 2B).

BLASTX analyses indicated that PIGPI is most similar to sequences of plant origin. However, when partial sequences were used, the comparison revealed that certain regions of the deduced protein share a higher percentage of identical amino acid residues (up to 67%) with a GPI homolog from *T. gondii*. Among 100 of the sequences most similar to PIGPI there are several hits corresponding to GPI proteins from bacterial and vertebrate organisms. On the other hand, at the nucleotide level, the highest similarity observed is with genes from bacterial and protozoan sources. Moreover, very low nucleotide sequence similarity was observed with GPI genes from plant or fungal origin.

Although the phylogenetic classification of *P. infestans* and its placement as member of the Kingdom Stramenopila are no longer a matter of debate (Alexopoulos et al., 1996), its origin is, nonetheless, a very intriguing subject. In this study, we have used four different methods, including distance-based and character-state (maximum parsimony) approaches, to reconstruct the evolutionary history of *P. infestans*, based on GPI nucleotide and protein sequences. Tree topologies were almost identical, regardless of the model used in distance calculation or the method used to reconstruct phylogeny (data not shown). For accuracy purposes, we decided to base this part of our discussion on the results obtained with the distance-based Neighbor-joining method, using protein sequences. The most noticeable attribute of the phylogenetic tree is the presence of a cluster containing the majority of *P. infestans* isolates and three other *Phytophthora* spp., a topology supported by 910 bootstrap replications. All sequences obtained from the isolates carrying the GPI profiles 100/111/122 (FL-01-3) and 100/122 (US97-001) were placed within the same subgroup in the major cluster; however, no correlation appears to exist between the GPI profiles and the general distribution of all isolates within the phylogenetic tree.

In our study, all phylogenetic trees resolved the sequences obtained from *P. cactorum*, a result supported by a bootstrap value of 86%, but failed to clearly separate both sequences from the *P. infestans* group. This outcome differs from previous molecular phylogenetic analyses showing that *P. cactorum* could be distinguished from *P. infestans* using ITS sequences (Cooke et al., 2000). Conversely, the phylogenetic placement of *P. mirabilis*, as sharing the most recent ancestor with members of the *P. infestans* group, is consistent with previous reports indicating the close relatedness between the two species (Goodwin et al., 1999). However, in our case, neither phylogenetic tree had a bootstrap value that validates the definitive existence of *P. mirabilis* as independent taxa (Fig. 5A).

Two examples of considerable molecular evolutionary divergence are observed within the GPI phylogeny. In the case of the Welsh isolate 98-139, nucleotide changes in *PIGPI* (1.4%) have led to the development of two phylogenetically distinct protein sequences located in the major and the minor clusters, respectively. Furthermore, the GPI sequences from the metalaxyl-resistant isolate of *P. erythroseptica* and the Ecuadorian isolate EC3215 of *P. infestans* were equally distributed in both clusters of the phylogenetic tree (Fig. 5A), a result that is in conflict with the ITS-based phylogeny, which clearly placed *P. erythroseptica* and *P. infestans* in separate clades (Cooke et al., 2000). Two plausible

alternatives could help explain this outcome. It is possible that bilateral gene flow has given rise to both types of alleles and further evolution led to very divergent protein sequences. Interspecific hybridization in the genus *Phytophthora* has been documented and it has been suggested that along with geographic radiation and host specialization, it may have played a role in the evolution of modern *Phytophthora* species (Cooke et al., 2000). In addition, lateral transfer of *GPI* already has been reported in other organisms (Katz, 1996; Nowitzki et al., 1998). On the other hand, misidentification of one of the samples could have led to the seemingly mixed origin of the sequences. In the absence of evidence supporting the hypothesis of potential gene transfer between *P. infestans* isolates, and even in the instance of a misidentification case, our results would suggest that two distinct classes of alleles have developed recently in the evolutionary scale.

We analyzed the phylogenetic relationships between the *P. infestans* consensus GPI and other prokaryotic and eukaryotic sequences available in the public databases and found that, thus far, PIGPI appears to be most closely related to genes from plant origin (Fig. 5B). Furthermore, the *P. infestans* consensus GPI and plant sequences form a subgroup clearly distinct from the cluster constituted by the remaining sequences analyzed in our study, which encompassed all major phylogenetic groups, as confirmed by a bootstrap value of 100%. This subgroup shares with *T. gondii* the most recent ancestor (Fig. 5B). It should be noted, however, that the absence of useful Stramenopile sequences in the public databases (only one short EST from the Stramenopile organism *Plasmopara halstedii* has been reported at the moment of writing this discussion, and because of its length, ~455 bp, it was not included in this study) is a limitation for a thorough, Kingdom-wide, phylogenetic analysis. Therefore, it is conceivable that the current phylogenetic status of PIGPI might change as soon as more sequences from the Kingdoms Stramenopila and Protista become available. In conclusion, our results from the GPI coding-region and protein sequence analyses reaffirm the phylogenetic status of *P. infestans*.

Although our investigation does not allow us to establish a direct correlation between an isolate's GPI profile and its virulence (nor was it intended to do so), we cannot disregard that the highly virulent Florida isolate (a member of race US-8) contains multiple alleles of the *GPI* gene, which show a significant amount of sequence variation. For example, Florida isolate's type II allele, in addition to its increased transcriptional activity (as compared to type-I allele's expression level), showed one of the highest ratios of non-synonymous versus synonymous substitutions, suggesting that this gene might be selectively advantageous. Further research is needed to clearly demonstrate whether such evolutionary changes are somehow associated with this isolate's virulence or, perhaps, play a role in *P. infestans* pathogenicity as it was shown in the case of *X. campestris* (Tung and Kuo, 1999).

Acknowledgments

We greatly appreciate the technical support from Ms. Theresa Henderson and Mr. Brian Moravec. We express our gratitude to Dr. Bill Fry, Dr. Paul Tooley, and two anonymous reviewers for critical reading of the manuscript and excellent suggestions. We also thank Dr. Ken Deahl for providing us with isolates of *P. infestans*.

References

Achari, A., Marshall, S.E., Muirhead, H., Palmieri, R.H., Noltmann, E.A., 1981. Glucose-6-phosphate isomerase. Philos. Trans. R. Soc. Lond. B Biol. Sci. 293, 145–157.

Alexopoulos, C.J., Mims, C.W., Blackwell, M., 1996. Introductory Mycology, fourth ed. Wiley, New York. p. 868.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Blom, N., Gammeltoft, S., Brunak, S., 1999. Sequence- and structure-based prediction of eukaryotic protein phosphorylation sites. J. Mol. Biol. 294, 1351–1362.

Chen, J., Roxby, R., 1996. Characterization of a *Phytophthora infestans* gene involved in vesicle transport. Gene 181, 89–94.

Chen, Y., Roxby, R., 1997. Identification of a functional CT-element in the *Phytophthora infestans piypt1* gene promoter. Gene 198, 159–164.

Church, G.M., Gilbert, W., 1984. Genomic sequencing. Proc. Natl. Acad. Sci. USA 81, 1991–1995.

Cooke, D.E.L., Drenth, A., Duncan, J.M., Wagels, G., Brasier, C.M., 2000. A molecular phylogeny of *Phytophthora* and related oomycetes. Fungal Genet. Biol. 30, 17–32.

Daayf, F., Platt, H.W., 2000. Changes in metalaxyl resistance among glucose phosphate isomerase genotypes of *Phytophthora infestans* in Canada during 1997 and 1998. Am. J. Potato Res. 77, 311–318.

Dahlhoff, E.P., Rank, N.E., 2000. Functional and physiological consequences of genetic variation at phosphoglucose isomerase: heat shock protein expression is related to enzyme genotype in a montane beetle. Proc. Natl. Acad. Sci. USA 97, 10056–10061.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Froman, B.E., Tait, R.C., Gottlieb, L.D., 1989. Isolation and characterization of the phosphoglucose isomerase gene from *Escherichia coli*. Mol. Gen. Genet. 217, 126–131.

Fry, W.E., Goodwin, S.B., 1997. Re-emergence of potato and tomato late blight in the United States. Plant Dis. 81, 1349–1357.

Garelik, G., 2002. Taking the bite out of potato blight. Science 298, 1702–1704.

Goodwin, S.B., Cohen, B.A., Deahl, K.L., Fry, W.E., 1994. Migration from Northern Mexico as the probable cause of recent genetic changes in populations of *Phytophthora infestans* in the United States and Canada. Phytopathology 84, 553–558.

Goodwin, S.B., Legard, D.E., Smart, C.D., Levy, M., Fry, W.E., 1999. Gene flow analysis of molecular markers confirms that *Phytophthora mirabilis* and *P. infestans* are separate species. Mycologia 91, 796–810.

Goodwin, S.B., Schneider, R.E., Fry, W.E., 1995a. Use of cellulose-acetate electrophoresis for rapid identification of allozyme genotypes of *Phytophthora infestans*. Plant Dis. 79, 1181–1185.

Goodwin, S.B., Sujkowski, L.S., Dyer, A.T., Fry, B.A., Fry, W.E., 1995b. Direct detection of gene flow and probable sexual reproduction of *Phytophthora infestans* in Northern North America. Phytopathology 85, 473–479.

Green, J.B., Wright, A.P., Cheung, W.Y., Lancashire, W.E., Hartley, B.S., 1988. The structure and regulation of phosphoglucose isomerase in *Saccharomyces cerevisiae*. Mol. Gen. Genet. 215, 100–106.

Grunwald, N.J., Flier, W.G., Sturbaum, A.K., Garay-Serrano, E., van den Bosch, T.B.M., Smart, C.D., Matuszak, J.M., Lozoya-Saldana, H., Turkensteen, L.J., Fry, W.E., 2001. Population structure of *Phytophthora infestans* in the Toluca valley region of central Mexico. Phytopathology 91, 882–890.

Hansen, J., Lund, O., Tolstrup, N., Gooley, A., Williams, K., Brunak, S., 1998. NetOglyc: prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. Glycoconj. J. 15, 115–130.

Kaslow, D.C., Hill, S., 1990. Cloning metabolic pathway genes by complementation in *Escherichia coli*—Isolation and expression of *Plasmodium falciparum* glucose phosphate isomerase. J. Biol. Chem. 265, 12337–12341.

Katz, L.A., 1996. Transkingdom transfer of the phosphoglucose isomerase gene. J. Mol. Evol. 43, 453–459.

Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.

Liu, Y.G., Whittier, R.F., 1995. Thermal asymmetric interlaced PCR—automatable amplification and sequencing of insert end fragments from P1 and Yac clones for chromosome walking. Genomics 25, 674–681.

McLeod, A., Smart, C.D., Fry, W.E., 2003. Characterization of 1,3-beta-glucanase and 1,3;1,4-beta-glucanase genes from Phytophthora infestans. Fungal Genet. Biol. 38, 250–263.

Nakai, K., Kanehisa, M., 1992. A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics 14, 897–911.

Nowitzki, U., Flechner, A., Kellermann, J., Hasegawa, M., Schnarrenberger, C., Martin, W., 1998. Eubacterial origin of nuclear genes for chloroplast and cytosolic glucose-6-phosphate isomerase from spinach: sampling eubacterial gene diversity in eukaryotic chromosomes through symbiosis. Gene 214, 205–213.

Pedersen, A.G., Nielsen, H., 1997. Neural network prediction of translation initiation sites in eukaryotes: perspectives for EST and genome analysis. Proc. Int. Conf. Intel. Syst. Mol. Biol. 5, 226–233.

Pieterse, C.M., van West, P., Verbakel, H.M., Brasse, P.W., van den Berg-Velthuis, G.C., Govers, F., 1994. Structure and genomic organization of the *ipiB* and *ipiO* gene clusters of *Phytophthora infestans*. Gene 138, 67–77.

Pieterse, C.M., van't Klooster, J., van den Berg-Velthuis, G.C., Govers, F., 1995. NiaA, the structural nitrate reductase gene of *Phytophthora infestans*: isolation, characterization and expression analysis in *Aspergillus nidulans*. Curr. Genet. 27, 359–366.

Raymond, J., Zhaxybayeva, O., Gogarten, J.P., Gerdes, S.Y., Blankenship, R.E., 2002. Whole-genome analysis of photosynthetic prokaryotes. Science 298, 1616–1620.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Thomas, B.R., Ford, V.S., Pichersky, E., Gottlieb, L.D., 1993. Molecular characterization of duplicate cytosolic phosphoglucose isomerase genes in *Clarkia* and comparison to the single gene in *Arabidopsis*. Genetics 135, 895–905.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Tooley, P.W., Fry, W.E., Gonzalez, M.J.V., 1985. Isozyme characterization of sexual and asexual *Phytophthora infestans* populations. J. Hered. 76, 431–435.

Tung, S.Y., Kuo, T.T., 1999. Requirement for phosphoglucose isomerase of *Xanthomonas campestris* in pathogenesis of citrus canker. Appl. Environ. Microbiol. 65, 5564–5570.