Plant Pathology Bulletin 12:255-268, 2003

Comparison of glyceraldehyde-3-phosphate dehydrogenase genes in *Phaeosphaeria nodorum* and *P. avenaria* species

P. P. Ueng [1,5], E. Reszka [2], K. R. Chung [3], E. Arseniuk [2] and J. M. Krupinsky [4]

1 Molecular Plant Pathology Laboratory, U.S. Department of Agriculture, ARS, Beltsville, MD 20705, USA
2 Department of Plant Pathology, Plant Breeding and Acclimatization Institute, Radzikow, Poland
3 Citrus Research & Education Center, University of Florida, Lake Alfred, FL 33850, USA
4 Northern Great Plains Research Laboratory, U.S. Department of Agriculture, ARS, Mandan, ND 58554, USA
5 Corresponding author, E-mail: uengp@ba.ars.usda.gov; Fax: 301-504-5449
Accepted for publication: November 10, 2003

ABSTRACT

Ueng, P. P. [1,5], Reszka, E. [2], Chung, K. R. [3], Arseniuk, E. [2] and Krupinsky, J. M. [4] 2003. Comparison of glyceraldehyde-3-phosphate dehydrogenase genes in *Phaeosphaeria nodorum* and *P. avenaria* species. Plant Pathol. Bull. 12:255-268.

Stagonospora leaf blotch of cereals is a complex disease caused by several *Phaeosphaeria* fungal species. Genetic diversity among glyceraldehyde-3-phosphate dehydrogenase (*gpd*) genes of *Phaeosphaeria nodorum*, *P. avenaria* f. sp. *triticea* (P.a.t.) and *P. avenaria* f. sp. *avenaria* (P.a.a.), three causal agents of stagonospora leaf blotch diseases in cereals, was evaluated in this study. The PCR-amplified *gpd* gene from these cereal *Phaeosphaeria* species contained 4 introns. The *gpd* gene in wheat-biotype *P. nodorum* isolates ranged in size from 1,253-1,255bp. Sequence variations in intron 4 were used to divide the wheat-biotype *P. nodorum* isolates into two subgroups. These two subgroups were equally represented in a survey of field wheat-biotype *P. nodorum* isolates. Except Pat2, the *gpd* genes in Pat1 and Pat3 of the P.a.t. group, the P.a.a. group and the barley-biotype *P. nodorum* were 1,251bp in length and phylogenetically closely related. As the nucleotide changes in *gpd* genes occurred mostly in intron regions and the third codons of amino acid coding triplets, the deduced protein sequences in three cereal *Phaeosphaeria* species were identical, except for Pat2. The *gpd*-encoded protein in Pat2 isolates contained two amino acid substitutions. The deduced GPD proteins of three cereal *Phaeosphaeria* species are closely related to the GPD protein that has been described in a maize pathogen, *Cochliobolus heterostrophus*. Based on phylogenetic analysis, the possible evolution of these cereal *Phaeosphaeria* species in relation to other ascomycetes is discussed.

Key words: glyceraldehyde-3-phosphate dehydrogenase gene, *Phaeosphaeria* spp., Stagonospora leaf blotch, wheat

INTRODUCTION

Glyceraldehyde-3-phosphate dehydrogenase (GPD) (EC 1.2.1.12) is an essential tetrameric and NAD-binding enzyme involved in glycolysis and glyconeogenesis in living organisms. Both the messenger RNA and its encoded GPD protein are two of the most abundant molecules in bakers' yeast cells [12,18]. The high levels of *gpd* gene expression in cells has led to the utilization of its strong transcriptional promoter to express heterologous genes and the development of transformation systems for many fungi [4,11,15,35,37]. Sequence diversity of partial *gpd* genes in combination with other genes has been used in phylogenetic studies in fungal pathogens. In a study of *Cochliobolus* (telemorph) and *Curvularia* and *Bipolaris* (anamorphs) species in Pleosporaceae, the *Cochliobolus* species with high virulence to cereals and common grasses clustered as a group [3]. Many *Stemphylium* spp. with unknown sexual states were grouped with those species belonging to *Pleospora* and the classification of several *Stemphylium* species pathogenic to alfalfa was thus clarified [5]. In addition, host specialization and geographic separation were proposed to be the evolutionary force in diversity among the tree pathogen *Gremmeniella* spp. [9]. Since the change rate of GPD proteins is one of the slowest known, the molecular data from various living organisms have been used for evolutionary studies [6,21,23,27].

Septoria and stagonospora leaf spot of cereals is a disease complex caused by a number of fungi that are not always closely related [24,28]. The *Septoria* and *Stagonospora*

species important on cereals and having great impact on agriculture are placed in the family Loculoascomycetes, with known teleomorphs in *Mycosphaerella* (anamorph *Septoria*) and *Phaeosphaeria* (anamorph *Stagonospora*) [7]. Four species, *Mycosphaerella graminicola* (anamorph *Septoria tritici*), *Septoria passerinii*, *Phaeosphaeria nodorum* and *P. avenaria*, are classified mainly on morphology and their host specificity. Differentiation of three cereal *Phaeosphaeria* species has recently been supplemented with important molecular data [7,29,30,31,32,33]. Based on restriction fragment length polymorphism (RFLP) fingerprints, rDNA internal transcribed spacer (ITS) sequences and mating type gene conserved region sequences, the genetic relationships among the wheat-biotype and barley-biotype in *P. nodorum*, three genetically distinct groups (Pat1, Pat2 and Pat3) in *P. avenaria* f. sp. triticea, and *P. avenaria* f. sp. avenaria have been defined [30,31,32,33]. The goal of this study was to examine the sequence diversity among *gpd* genes of three cereal *Phaeosphaeria* species to support species designations and their evolution.

MATERIALS AND METHODS

Fungal isolates and DNA isolation

Eight wheat-biotype *P. nodorum*, five barley-biotype *P. nodorum*, nine *P. avenaria* f. sp. *triticea* (P.a.t.), including six Pat1, two Pat2 and one Pat3, and four *P. avenaria* f. sp. *avenaria* (P.a.a.) isolates were used for *gpd* gene sequence analysis (Table 1). Most procedures for growing fungal cultures in a liquid medium and the isolation of genomic DNA were described by Ueng et al. [29].

gpd gene amplification, sequencing and data analysis

Three sets of oligonucleotide primers, designed from a *gpd* genomic DNA sequence of the *P. nodorum* isolate BSm300 (GenBank accession number AJ271155), for PCR amplification of the same genes from other cereal *Phaeosphaeria* species. Individual fragments of 1,026, 1,112 and 1,166 bp in length could be produced from each respective primer set of 14A/14B (nucleotide #1584-1605/ #2609-2590) (5'-AAACACGCGCCCGAGGTAGCAT/5'-GCGAGGGGAGCCAAGCAGTT), 4A/4B (nucleotide #2015-2035/#3127-3107) (5'-CAACGGCTTCGGTCGCATTGG/ 5'-CTCGTTGTCGTACCAGGAGAC) and 12A/12C (nucleotide #2391-2410/#3556-3535) (5'-CGAGACCGGCGCATACTACG/ 5'-CTCACGCGACTCCAGGTTTTCT). These three fragments should overlap and represent the full-length *gpd* gene-coding region. PCR amplification was performed in a 50 μl reaction mixture containing 1x reaction buffer (50mM KCl, 10mM Tris-HCl, pH9.0 at 25℃, with 0.1% Triton X-100), 1.25mM $MgCl_2$, 0.2mM dNTPs, 2 μM each primers, 80ng genomic DNA and 1.0unit of Taq DNA polymerase (Promega, Madison, WI). Reaction parameters were: denaturing at 94℃ for 3 min, amplifying by 40 cycles of 94℃ for 20 s, 55℃ for 30 s and 72℃ for 1 min, and then extending at 72℃ for 10 min. The PCR products were resolved by 0.8% agarose gel electrophoresis, eluted from gel blocks with a unidirectional electroelutor (Model UEA, IBI, New Haven, CT), concentrated with 1-butanol, extracted with 1X volume phenol/chloroform mixture (v/v 1:1), and precipitated with an equal volume of 2- propanol.

For further sequencing in the intron 4 region of *gpd* gene in wheat-biotype *P. nodorum*, genomic DNA from 18 New York isolates (237, 238, 240, 242, 244, 403, 405, 406, 407, 408, 409, 410, 411, 412, 414, 417, 419, 420) used in a previous study [26] and 8 Indiana isolates (IN15, IN37, IN38, IN39, IN43, IN46, IN9011, S-97-2) were used as templates for PCR amplification. With the same PCR reaction mixture and parameters mentioned above, ca. 400 bp size fragments were produced with the specific primer set, 42A/4B (nucleotide #2726-2747/#3127-3107) (5'-GCACTGCGGCCCAGAACATCAT/5'-CTCGTTGTCGTACCAGGAGAC). The PCR products were purified by gel electro- elution as mentioned above.

PCR-amplified DNA fragments were directly sequenced with an automatic DNA sequencer (ABI PRISM 3700, Applied Biosystems, Foster City, CA) using the same primers as used for PCR amplification. Reactions and analyses of PCR-sequencing followed the procedures described in a published protocol (ABI PRISM BigDye Terminators v3.0 Cycle Sequencing Kit). The consensus nucleotide sequence of the *gpd* gene for each isolate was determined by overlapping multiple sequence data.

Twenty-five representative full-length *gpd* gene nucleotide and their deduced peptide sequences from ascomycetes deposited in GenBank (accession #'s are listed in Table 2 and Fig. 4), and, together with those of three cereal *Phaeosphaeria* spp. (accession #'s are listed in Table 1), were used for Phylogenetic Analysis Using Parsimony (PAUP) analysis. GPD peptide sequences from twelve basidiomycetes and two zygomycetes (accession #'s are listed in Figs. 3 and 4) were also compared. The 331-338 amino acid sequences and their corresponding nucleotides (996-1,455 bp) in ascomycetes, and 335- 343 amino acids and 1014-1717 bp nucleotide sequences in basidiomycetes and zygomycetes were aligned using the PILEUP program in the Wisconsin Sequence Analysis package of the Genetics Computer Group (Madison, Wisconsin). Maximum parsimony trees were inferred by using the heuristic search option of bootstrap analysis in PAUPSEARCH program. The PAUPDISPLAY program drew trees to indicate phylogenetic relationship.

RESULTS

***gpd* gene structure and sequence analyses.** Using the primers described in the Materials and Methods, a 1,251 bp fragment of the *gpd* gene was obtained from P.a.t., P.a.a. and

Table 1. Isolates of *Phaeosphaeria* species used in genetic analysis

Species	Original host	Geographic location	GenBank accession number
Phaeosphaeria nodorum (wheat-biotype)			
9074	Wheat	Gallatin, MT	AY364462
9076	Wheat	Richland, MT	
8408	Wheat	Mandan, ND	
9506	Barley	Mandan, ND	
S-74-20A (ATCC200806)	Wheat	Griffin, GA	AY364461
sn26-1	Wheat	Rzeszow, Poland	
sn27-1	Wheat	Sieradz, Poland	
sn37-1	Wheat	Szelejewo, Poland	AY364460
Phaeosphaeria nodorum (barley-biotype)			
S-82-13 (ATCC200805)	Barley	Senoia, GA	AY364464
S-83-2 (ATCC200841)	Barley	Tifton, GA	
S-83-7	Barley	Holland, VA	
S-84-2	Barley	Moultrie, GA	
S-92-7	Barley	Raleigh, NC	
Phaeosphaeria avenaria f. sp. *triticea* (Pat1)			
10052-2	Wheat	Langdon, ND	
12618	Wheat	Dickinson, ND	
12889	Wheat	Mandan, ND	
13061	Barley	Morton County, ND	
Sat24-1	Wheat	Warmi?sko-Mazurskie, Poland	AY364463
Sat28-1	Barley	Mazowieckie, Poland	
Phaeosphaeria avenaria f. sp. *triticea* (Pat2)			
ATCC26370	Foxtail barley	Minnesota	AY364457
ATCC26377	Foxtail barley	Minnesota	
Phaeosphaeria avenaria f. sp. *triticea* (Pat3)			
S-81-W10	Wheat	Washington	AY364456
Phaeosphaeria avenaria f. sp. *avenaria*			
ATCC12277	Oat	USA	AY364459
SAA001NY-85	Oat	New York	
SAT001NY-84	Wheat	New York	
SAT002NY-84	Wheat	New York	

barley-biotype *P. nodorum* isolates, with the exception of the Pat2 of P.a.t (Fig. 1). Two isolates representing pat2 ATCC26370 and ATCC26377 from foxtail barley (*Hordeum jubatum* L.) had a corresponding *gpd* gene of 1,252 bp in length (Figure 1). Size variations of *gpd* genes were found in wheat-biotype *P. nodorum* isolates. A *gpd* gene of 1,254 bp was detected in isolates 9074 and 9076, of 1,253 bp in isolates 8408, 9506, S-74-20A, sn26-1 and sn27-1, and of 1,255 bp in the isolate sn37-1 (Figure 1).

The *gpd* gene structures in these *Phaeosphaeria* species are identical. The gene contains 5 exons and 4 introns (Figure 1). The intron sizes were determined to be 55, 56-57, 79-80 and 47-49 bp, respectively. The intron-exon borders in these cereal *Phaeosphaeria gpd* genes are highly similar to known intorn-exon consensus sequences in filamentous fungi [1]. The consensus sequences (A/C)(C/T)AG for filamentous fungi were found at 3' splice sites in each intron (Figure 1). For the 5' splice site, the intron-exon junction for intron 1 was similar to those from higher eukaryotes (C/A)AG | GT(A/G)AG, and the other three were either yeast-like (N | GTA(T/C)GT in introns 2 and 4) or filamentous fungi-like (GTA(A/T/C)G in intron 3) [1]. In addition, a sequence motif, tgCT(A/G)ACX, important for RNA splicing within the introns of *Neurospora crassa* and yeast genes was also found in introns of *gpd* genes in these cereal *Phaeosphaeria* species (Figure 1) [17,20]. These sequences were 11, 9, 9 and 8 bp upstreams from 3' splicing sites of each respective intron. In general, most of the 5th introns in *gpd* genes in ascomycetes, which is intron 4 in these cereal *Phaeosphaeria* species, are AT-rich (Table 2).

Most nucleotide variation in the *gpd* genes occurred in the introns 2, 3 and 4 in these cereal *Phaeosphaeria* species. Furthermore, the variations in the intron 4 could divide wheat-biotype *P. nodorum* isolates into two separate groups. Group I had a nucleotide sequence consisting of a poly A plus

Table 2. Comparison of numbers and sizes of introns and their GC contents in glyceraldehyde-3- phosphate dehydrogenase genes of several ascomycetes

Species	Introns [1]					
	1	2	3	4	5	6
	▲	▲	■	▲	●	●
	V (5)	V (17)	A (21)	A (43)	G (271)	V (328)
Phaeosphaeria nodorum (wheat) groupI	55 (53)[2]	---	56 (48)	80 (58)	49 (29)	---
P. nodorum (wheat) groupII	55 (53)	---	56 (48)	80 (59)	48 (31)	---
P. nodorum (barley)	55 (55)	---	56 (48)	79 (54)	47 (36)	---
P. avenaria f. sp. *avenaria*	55 (53)	---	56 (50)	79 (56)	47 (34)	---
P. avenaria f. sp. *triticea* (Pat1)	55 (55)	---	56 (52)	79 (52)	47 (36)	---
P. avenaria f. sp. *triticea* (Pat2)	55 (51)	---	57 (47)	79 (53)	47 (38)	---
P. avenaria f. sp. *triticea* (Pat3)	55 (55)	---	56 (52)	79 (52)	47 (36)	---
Aspergillus oryzae (AF320304)	---	60 (**40**)	125 (54)	81 (48)	56 (**30**)	60 (47)
Acremonium chrysogenum (E02656)	---	---		217 (53)	65 (51)	---
Ajellomyces capsulatus (AF273703)	93 (**36**)	84 (51)	72 (49)	96 (**34**)	58 (**35**)	38 (**40**)
Blumeria graminis hordei (X99732)	---	---	---	51 (**33**)	46 (**26**)	---
Claviceps purpurea (X73282)	---	---	---	355 (55)	---	---
Coccidioides posadasii (AF288134)	69 (**35**)	63 (51)	79 (47)	93 (46)	73 (**30**)	54 (39)
Cochliobolus heterostrophus (X63516)	59 (56)	---	59 (49)	105 (52)	50 (**38**)	---
Cochliobolus lunatus (X58718)	59 (68)	---	57 (47)	120 (49)	53 (**34**)	---
Colletotrichum lindemuthianum (AF000310)	---	---	---	97 (46)	---	---
Cryphonectria parasitica (X53996)	---	---	---	150 (49)	72 (45)	---
Emericella nidulans (M19694)	64 (41)	56 (54)	109 (**40**)	58 (**40**)	54 (**37**)	51 (47)
Glomerella cingulata (M93427)	---	---	---	216 (54)	---	---
Monascus purpureus (Z68498)	---	65 (46)	64 (53)	72 (53)	58 (43)	---
Neurospora crassa (U67457)	---	---	---	74 (47)	---	---
Paracoccidioides brasiliensis (AF396657)	---	48 (56)	68 (41)	92 (**36**)	72 (**38**)	---
Podospora anserina (X62824)	---	---	---	77 (59)	---	---
Sclerotinia sclerotiorum (AF417110)	---	---	---	65 (51)	54 (**22**)	---
Sordaria macrospora (AJ313527)	---	---	---	76 (50)	---	---
Tolypocladium inflatum (A40404)	---	---	---	417 (57)	---	---

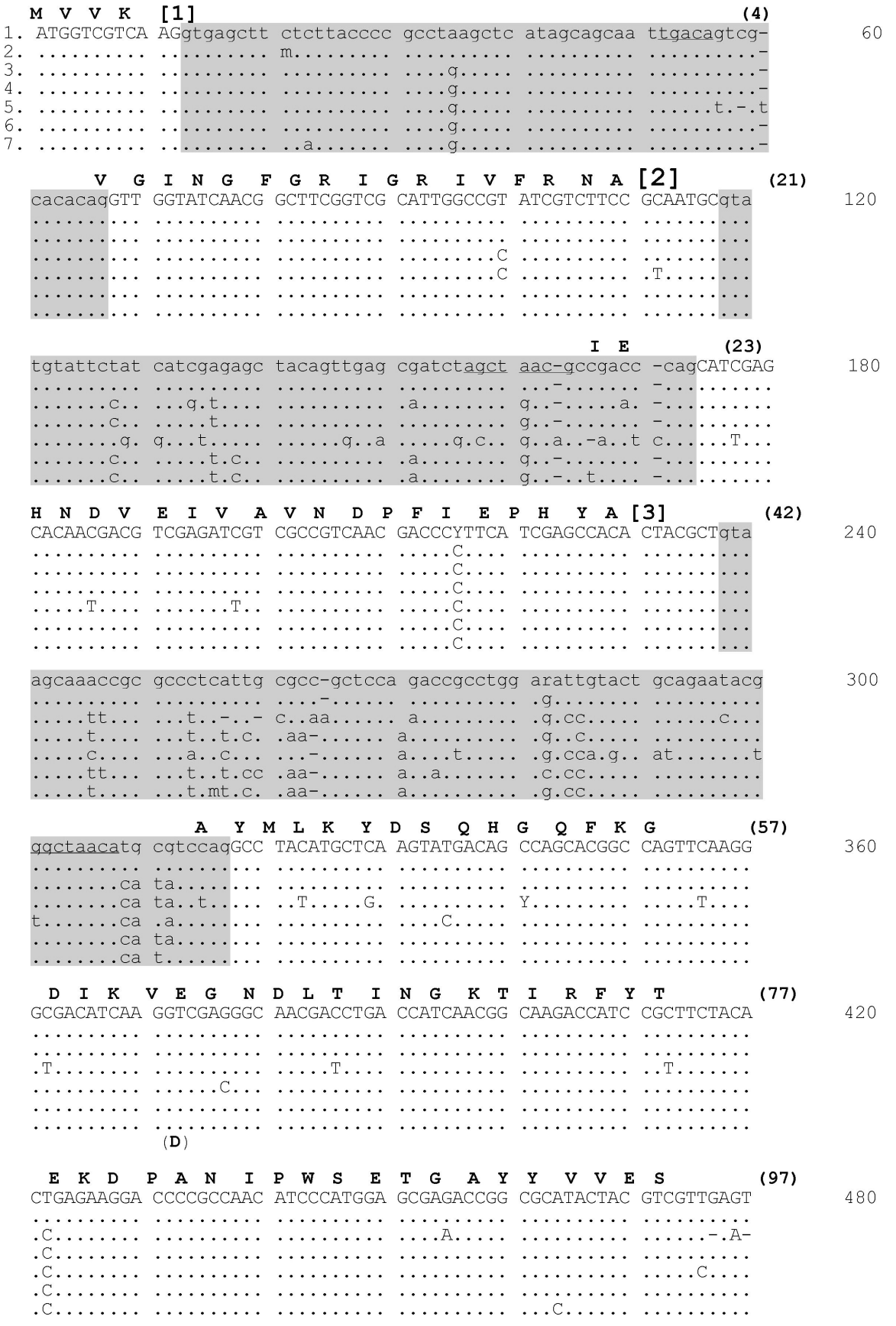
[1] Six introns indicated by amino acids and their positions (in parentheses) in deduced glyceraldehyde-3-phosphate dehydrogenase peptides of *Phasosphaeria nodorum*. Symbols represent the position within the codon where the intron occurs: ▲ before the first base; ● before the second base; and ■ before the third base of the codon.

[2] Number of nucleotides (bp) presents in each intron. The percentage of GC contents is in parentheses. The AT rich sequences are in bold. '---' indicates no intron at particular position.

CAC and Group II had a sequence of AAGAAACA plus CAA in the middle portion of intron 4 (Figure 1). In Group I, a stretch of 8, 9 and 10 poly As was found in isolates 9506, 9074/9076/BSm300 and sn37-1, respectively. Group II included four other wheat-biotype *P. nodorum* isolates 8408, S-74-20A, sn26-1 and sn27-1.

In the partial *gpd* gene fragments amplified by 42A/4B primers, the intron 4 sequences in 10 wheat-biotype *P. nodorum* isolates collected from fields in New York and Indiana were similar to those of Group I, and 14 others were determined to belong to Group IIA (Figure 2). In Group I, five isolates, IN15, 405, 406, 417, 419, had a stretch of 9 A residues (Group IA) and the other five, IN39, 244, 407, 414, 420, had 10 A residues (Group IB). As compared to Group IIA *P. nodorum* isolates, two nucleotide substitutions (#238 and #240) were found in intron 4 of two isolates, IN38 and 410 (Figure 2, Group IIB).

When compared the *gpd* gene of wheat-biotype *P. nodorum* Group I, the number of nucleotide substitutions in the wheat-biotype *P. nodorum* Group II, barley-biotype *P. nodorum*, P.a.a., and Pat1, Pat2 and Pat3 of P.a.t. were 8, 57, 57, 54, 88 and 58 bp, respectively. The nucleotide sequence identity with wheat-biotype *P. nodorum* Group I was in the range of 93% for Pat2 to 99.4% for wheat-biotype *P. nodorum* Group II. Sequence variation was also found in *gpd* gene coding regions in intra- and inter-cereal *Phaeosphaeria* species. Since the nucleotide changes only occurred in the third codons of amino acid coding triplets, they do not affect the amino acid composition. The deduced GPD enzymes containing 337 amino acids were identical in all three cereal *Phaeosphaeria* species, except for the two isolates (ATCC26370 and ATCC26377) from foxtail barley in Pat2 of P.a.t. Nucleotide substitutions at #377 (from 'G' to 'C') and at #972 (from 'G' to 'A') in the *gpd* gene altered the peptide


M V V K [1] (4)
1. ATGGTCGTCA AGgtgagctt ctcttacccc gcctaagctc atagcagcaa ttgacagtcg- 60
2. .... m......... ....-
3. ....g..... ....-
4. ....g..... ....-
5. ....g..... ....t.-.t
6. ....g..... ....-
7. ....a....... ....g..... ....-
V G I N G F G R I G R I V F R N A [2] (21)
cacacagGTT GGTATCAACG GCTTCGGTCG CATTGGCCGT ATCGTCTTCC GCAATGCgta 120
I E (23)
tgtattctat catcgagagc tacagttgag cgatctagct aac-gccgacc -cagCATCGAG 180
H N D V E I V A V N D P F I E P H Y A [3] (42)
CACAACGACG TCGAGATCGT CGCCGTCAAC GACCCYTTCA TCGAGCCACA CTACGCTgta 240
agcaaaccgc gccctcattg cgcc-gctcca gaccgcctgg arattgtact gcagaatacg 300
A Y M L K Y D S Q H G Q F K G (57)
ggctaacatg cgtccagGCC TACATGCTCA AGTATGACAG CCAGCACGGC CAGTTCAAGG 360
D I K V E G N D L T I N G K T I R F Y T (77)
GCGACATCAA GGTCGAGGGC AACGACCTGA CCATCAACGG CAAGACCATC CGCTTCTACA 420
(D)
E K D P A N I P W S E T G A Y Y V V E S (97)
CTGAGAAGGA CCCCGCCAAC ATCCCATGGA GCGAGACCGG CGCATACTAC GTCGTTGAGT 480

```
  T  G  V  F  T  T  T  D  K  A  K  A  H  L  K  G  G  A  K  K       (117)
CCACCGGTGT CTTCACCACC ACCGACAAGG CCAAGGCCCA CTTGAAGGGT GGTGCTAAGA        540
.......... .......... .......... .......... .......... ..........
.T-.T..... ....-.G.-. .......... .......... .......... ..........
.......... .......... .......... .......... .......... ..A..C....
....T..... .......... .......... .......... .......... ..........
.......... ......G... .......... .......... .......... ..........
.......... ......A... .......... .......... .......... ..........

  V  V  I  S  A  P  S  A  D  A  P  M  F  V  M  G  V  N  N  E       (137)
AGGTCGTCAT CTCCGCTCCC TCGGCTGATG CKCCCATGTT CGTCATGGGT GTCAACAACG        600
.......... .......... .......... .G........ .......... ..........
.......... .......... .......... .G........ .......... ..........
.......... .......... .......... .G........ .......... ..........
.......... ...G...... .......... .G........ ...T...... ..........
.......... .......... .......... .G........ .......... ..........
.......... .......... .......... .T........ .......... ..........

  T  Y  T  K  D  I  E  V  L  S  N  A  S  C  T  T  N  C  L  A       (157)
AGACCTACAC CAAGGACATY GAGGTGCTCT CCAACGCCTC CTGCACAACC AACTGCTTGG        660
.......... .........C .......... .......... .......... ..........
.......... A........C .......... .......T.. .......... ..........
.......... A........C .......... .......... .......... ..........
.......... A........ .......... .......... G......... ..........
.......... A........C .......... .......T.. .......... ..........
.......... A........C .......... .......T.. .......... ..........

  P  L  A  K  V  I  H  D  K  F  T  I  I  E  G  L  M  T  T  V       (177)
CTCCCCTCGC CAAGGTCATC CACGACAAGT TCACCATCAT TGAGGGTTTG ATGACCACCG        720
.......... .......... .......... .......... .......... .........T
.......... .......... .......... .......... .......... ..........
....T..... .......... .......... .......... C.....C... .....T....
.......... .......... .......... .......... .......... .........T
.......... .......... .......... .......... .......... .........T

  H  S  Y  T  A  T  Q  K  V  V  D  G  P  S  A  K  D  W  R  G       (197)
TCCACTCCTA CACCGCTACC CAGAAAGTCG TTGACGGCCC TTCCGCCAAG GACTGGCGTG        780
.......... .......... .......... .......... .......... ..........
.......... ......C... .....G.... .......... G..T..T... ..........
.......... ......C... .....G.... .......... C..T..T... ..........
.......... .........A .....G.... ....T..... C..T..A... ..........
.......... ......C... ..A..G.... .......... C..T..T... ..........
.......... ......C... ..A..G.... .......... C..T..T... ..........

  G  R  T  A  A  Q  N  I  I  P  S  S  T  G  A  A  K  A  V  G       (217)
GTGGCCGCAC TGCGGCCCAG AACATCATTC CCAGCAGCAC TGGTGCCGCC AAGGCTGTCG        840
.......... .......... .......... .......... .......... ..........
.......... .......... .......... .......... .......... ..........
.......... .......... ........C. .......... .......... .........T
.......... .......... .......... .......... .......... ..........
.......... .......... .......... .......... .......... ..........
.......... .........R .......... .......... .......... ..........

  K  V  I  P  D  L  N  G  K  L  T  G  M  S  M  R  V  P  T  S       (237)
GCAAGGTCAT TCCTGACCTC AACGGCAAGC TCACCGGAAT GTCCATGCGT GTTCCCACCT        900
.......... .......... .......... .......... .......... ..........
....A..... .......... .......... .......... .......... ..C.......
.......... ......T... .......... .......... .......... ..........
.T........ .......... .......... .......... ...T...... ..........
.......... .......... .......... .......... .......... ..C.......
.......... .......... .......... .......... .......... ..........

  N  V  S  V  V  D  L  T  V  R  L  E  K  G  A  T  Y  D  E  I       (257)
CCAACGTTTC TGTGGTTGAC TTGACTGTCC GCCTCGAGAA GGGTGCTACC TACGATGAGA        960
.......... .......... .......... .......... .......... ..........
.......... C......... .......... .T........ .......... ..........
.......... C......... .......... .......... .......... ..........
.......... C......... .......... .......... .......G... .....C....
.......... C......... .......... .T........ ...C...... ..........
.......... C......... .......... .T........ .......... ..........
```

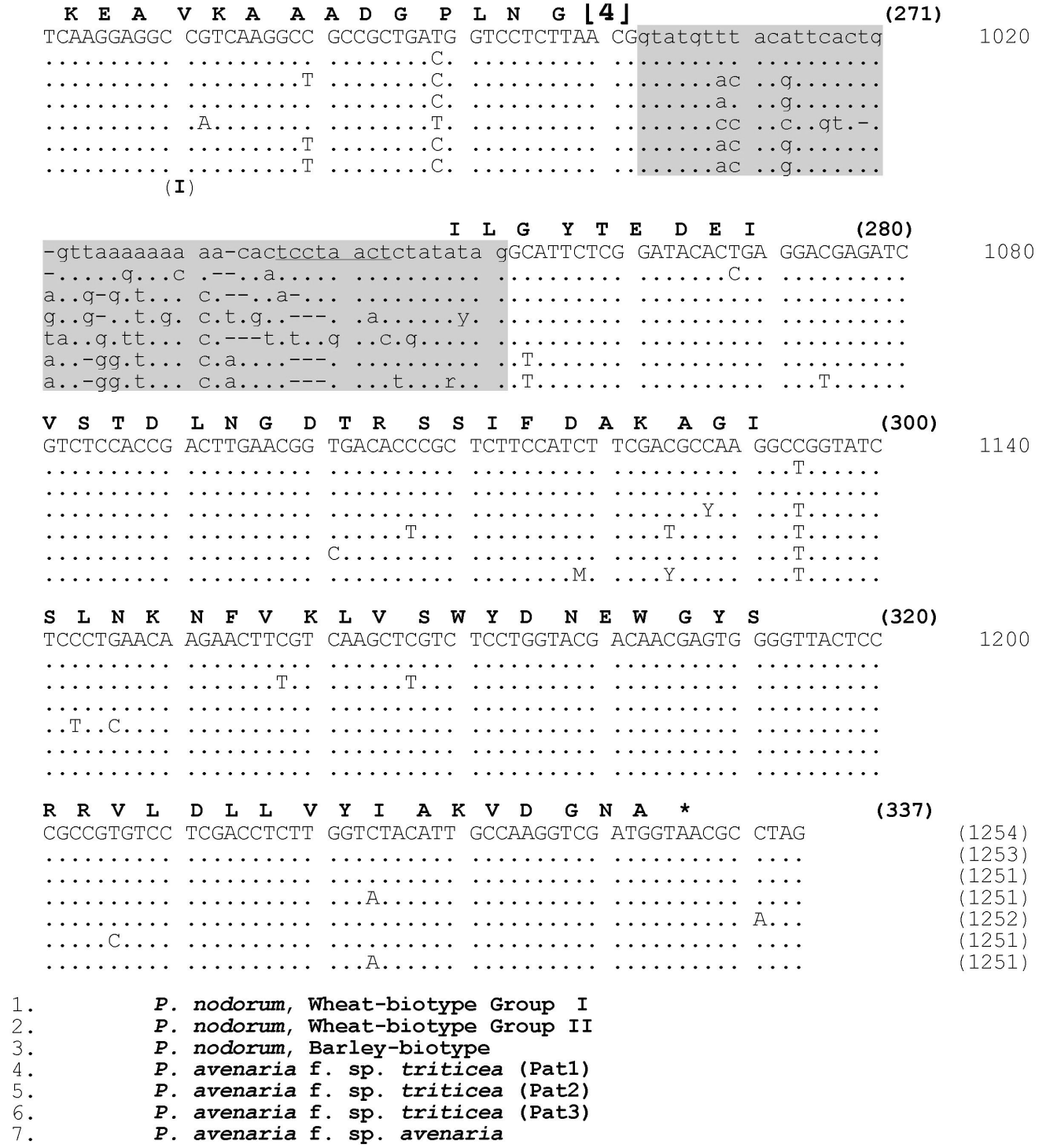


Fig. 1. Nucleotide sequence alignment of glyceraldehyde-3-phosphate dehydrogenase genes from cereal *Phaeosphaeria* species. Shaded regions [1]-[4] represent four introns, and the intron sequences are shown in lowercase letters. The nucleotides double underlined are essential RNA splicing sequences. Bold letters above the first sequence represent the deduced amino acids. Two amino acids in bold letters in parentheses below the seventh sequence indicate the amino acids substitutions in two Pat2 isolates. '.', Nucleotides are the same in reference to the first sequence. '-', Nucleotides are missing. The numbers in parentheses indicate the number of nucleotides in base pairs. The bold numbers in parentheses indicate the number of amino acids. Nucleotide substitutions found in each gene are represented by K = T or G; M = A or C; R = A or G; Y = T or C.

Fig. 2. Sequence comparison of partial glyceraldehyde-3-phosphate dehydrogenase gene fragments from wheat-biotype *Phaeosphaeria nodorum* field isolates. The fragments containing intron 4 (the shaded area and with lowercase letters) were amplified with the primer set 42A/4B. '.', Nucleotides are the same in reference to the first sequence. '-', Nucleotides are missing. The numbers in parentheses after each group indicate the number of isolates identified.

sequence from glutamic acid (E) to aspartic acid (D) and valine (V) to isoleucine (I), respectively, in Pat2 (Figure 1).

Phylogenetic analyses. In PAUP analyses, *gpd* gene nucleotide sequences of P.a.a, Pat1, Pat3 and barley-biotype *P. nodorum* were more closely related to each other than to Pat2 and wheat-biotype *P. nodorum* (Figure 3). Even though *gpd* genes of three cereal *Phaeosphaeria* species at nucleotide level have a low degree of similarity to and are phylogenetically distant from the other ascomycetes (Figure 3), their peptide sequences are closely related to *Cochliobolus heterostrophus*, a maize pathogen [3,36] (Figure 4). Based on the GPD peptide sequence analysis, fungi are initially divided into three distinct phylogenetic clades (Figure 4). The budding yeasts, most of the basidiomycetes, two zygomycetes and one ascomycete, *Magnaporthe grisea* (the rice blast pathogen), are closely related as a single group. *Schizosaccharomyces pombe* (fission yeast) and seven ascomycetes, including *Blumeria graminis* f. sp. *hordei* (barley powdery mildew), *Glomerella cingulata* (anthracnose) and *Sclerotinia sclerotiorum* (stem rot), cluster with two other basidiomycetes, *Ustilago maydis* (smut) and *Coprinopsis cinerea*. In contrast to *Colletotrichum lindemuthianum* (anthracnose), the other twelve ascomycetes, including the cereal *Phaeosphaeria* species, are grouped as the third clade in the evolutionary tree (Figure 4).

Differentiation of cereal *Phaeosphaeria* species by PCR amplification and enzymatic restrictions. Enzymatic restrictions of 4A/4B primers-amplified *gpd* gene fragments (1,112 bp) can be used to differentiate three cereal *Phaeosphaeria* species (Table 3). The endonuclease

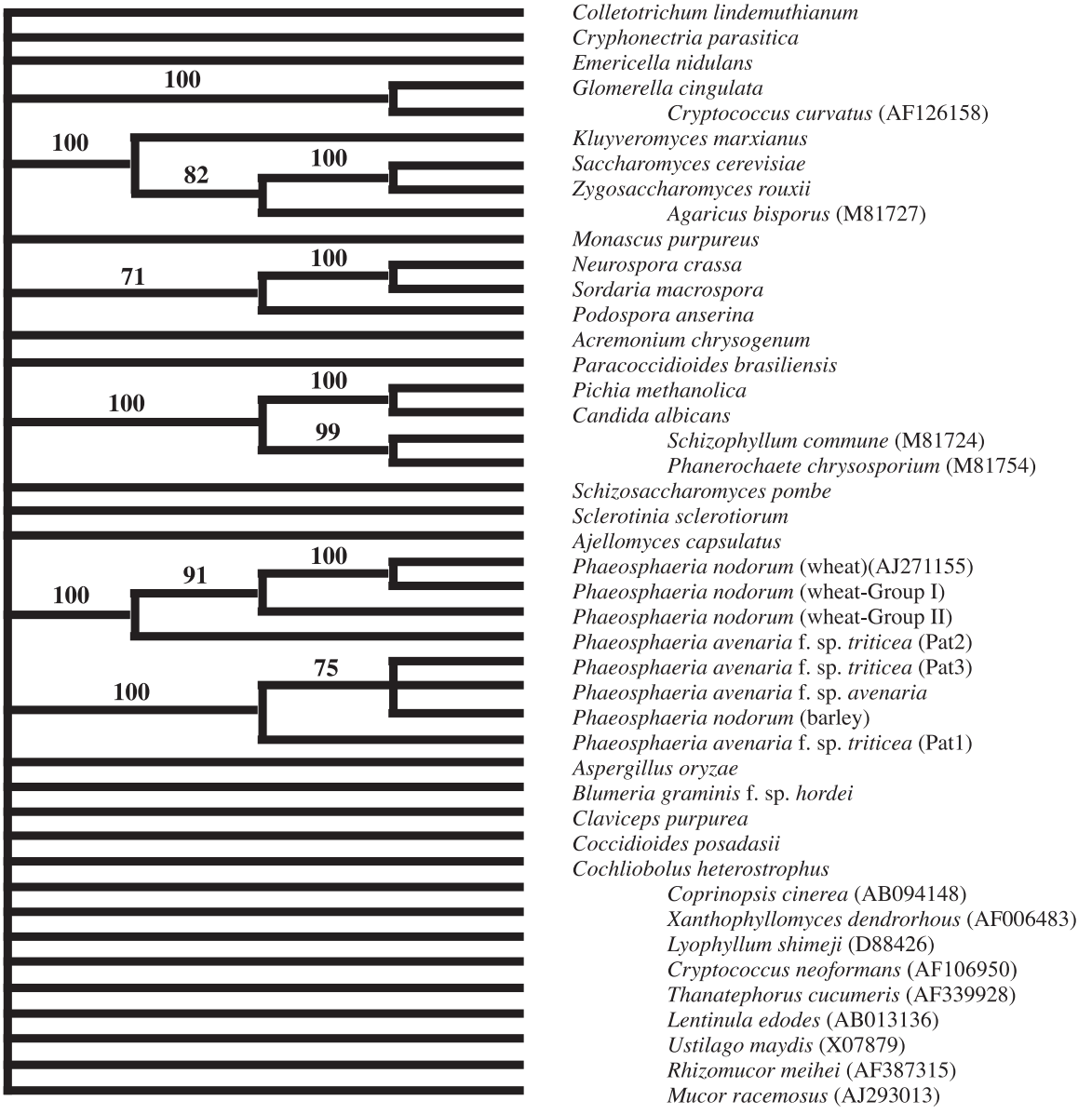


Fig. 3. Phylogenetic relationships based on nucleotide sequences of glyceraldehyde-3- phosphate dehydrogenase genes from ascomycetes, basidiomycetes and oomycetes. GenBank accession numbers are given in parentheses. Bootstrap values (with 100 replications) of the internal branches are indicated. Branches supported by less than 50% were collapsed to yield polytomies.

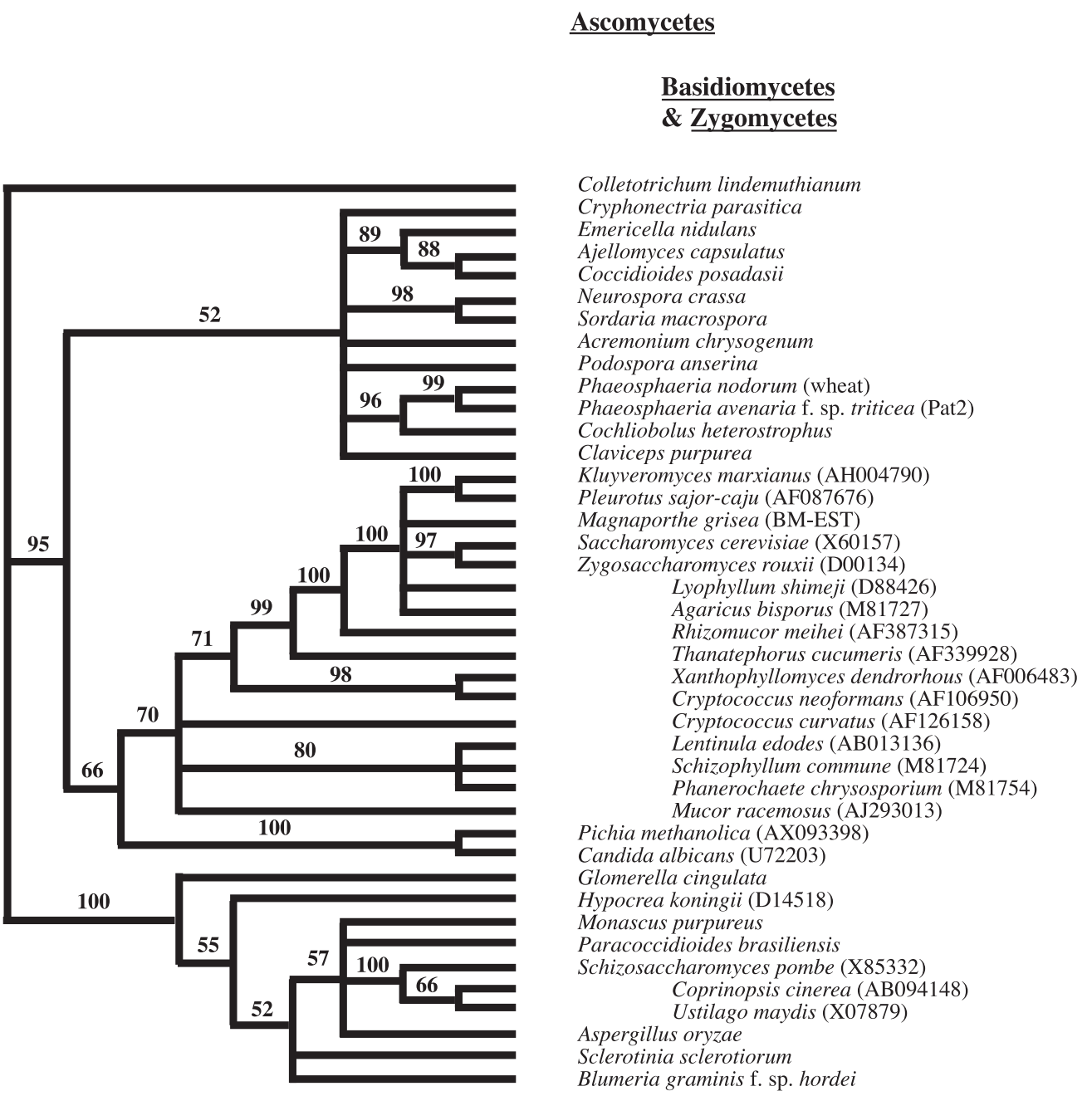


Fig. 4. Phylogenetic relationships based on deduced peptide sequences of glyceraldehyde-3-phosphate dehydrogenase genes from ascomycetes, basidiomycetes and oomycetes. GenBank accession numbers are given in parentheses. Two peptide sequences are either from m-RNA (*Hypocrea koningii*, D14518) or 41 expression sequence tags (*Magnaporthe grisea*, BM-EST). Bootstrap values (with 100 replications) of the internal branches are indicated. Branches supported by less than 50% were collapsed to yield polytomies.

Table 3. Comparison of endonuclease restriction maps of PCR-amplified glyceraldehyde-3-phosphate dehydrogenase gene products in *Phaeosphaeria* leaf pathogens

Endonuclease enzymes	*P. nodorum* (wheat-biotype)	*P. nodorum* (barley-biotype)	*P. a. t.* (Pat1) (ATCC26370)	*P.a.t.* (Pat2) (S-81-W10)	*P.a.t.* (Pat3)	*P.a.a.*
*Acc*I	-	403, 709	181, 931	297, 815	-	273, 839
*Age*I	407, 705	-	406, 706	-	407, 705	406, 706
*Dde*I	114, 121, 345, 532	47, 124, 231, 252, 458	124, 299, 689	123, 459, 530	47, 123, 231, 252, 459	47, 124, 231, 252, 458
*Hae*II	-	-	-	-	242, 870	386, 726
*Nar*I	-	-	-	-	-	383, 729
*Pst*I	217, 895	216, 896	216, 896	-	217, 895	216, 896
*Pvu*I	-	-	-	78, 1034	-	-
*Sal*I	-	-	-	296, 816	-	274, 838

Partial genes are amplified by primers 4A/4B. Fragment sizes are given in basepairs. '-' not cut by enzymatic restriction

restriction sites were deduced from the sequencing data and were experimentally demonstrated by enzymatic restrictions and agarose gel electrophoresis (*unpublished data*). Specific restrictions by individual *Age*I, *Nar*I, *Pst*I and *Pvu*I enzymes can identify barley-biotype *P. nodorum*, Pat2 and P.a.a. isolates. Distinct gel fingerprinting after digestions with respective *Acc*I, *Dde*I, *Hae*II and *Sal*I endonucleases can detect all three cereal *Phaeosphaeria* species (Table 3).

DISCUSSION

The *gpd* genes are discrete in many species studied, but are fused with other glycolytic genes in lower organisms. The *gpd* gene in Oomycota was reportedly fused with *tpi* gene encoding triose-phosphate isomerase [34]. Except in yeasts, one to six introns are present at conserved positions in the *gpd* genes of 20 ascomycetes (Table 2). All these intron-containing ascomycetes have an intron positioned at the #42/#43 amino acid junction in GPD peptides. In a recent survey, five identical intron positions reported to be conserved in nuclear genes of chloroplast/cytosol GPD enzymes in various diversified organisms [16]. This phenomenon supports the 'introns early' hypothesis and proposes that the introns were present in the earliest cells and facilitated the assembly of primordial genes by exon shuffling [10,16]. In all cereal *Phaeosphaeria* species, except two isolates belonging to Pat2, the deduced GPD proteins contained 337 identical amino acids. It appears that the *gpd* gene is well conserved during evolution of these organisms. In particular, the DNA sequences in intron 1 are very similar in these cereal *Phaeosphaeria* species (Figure 1). DNA sequences in the intron regions of the *GapC* (encoding GPD) [13], *Adh* (encoding alcohol dehydrogenase) [8], *CAD* (encoding cinnamyl alcohol dehydrogenase) [25] and *cdc2* (encoding cell cycle-regulated protein kinase) [19] genes in plants were used to investigate their intra- and inter-specific polymorphisms, evolutionary relationships and association with speciation. The intron length variation occurring in the *Adh* gene of *Brachyscome* was suggested to be relic of 5S DNA gene, short interspersed elements and fragments of transposons [8,14]. The origin of repeated A's in intron 4 of the *gpd* gene in Group I wheat-biotype *P. nodorum* isolates is unknown. Two different alleles found in *gpd* gene intron 4 in wheat-biotype *P. nodorum* isolates are either diverged from a species-specific ancestral sequence after their speciation or already existed in the ancestral species and then evolved to the present day.

Based on phylogenetic analysis of GPD peptide sequences, three cereal *Phaeosphaeria* species are closely related to *Cochliobolus* species (Figure 4). The deduced GPD peptides in *C. lunata* (X58718) and *C. heterostrophus* (X63516) have 92.3% and 92% in similarity to those in cereal *Phaeosphaeria* species [22,36]. When compared to the *gpd* gene coding sequence of *P. nodorum* wheat-biotype isolates, the similarity is 83.1% for *C. lunata* (X58718) and 82.3% *C. heterostrophus* (X63516). The *gpd* gene structures in both *Phaeosphaeria* and *Cochliobolus* species are similar; having 4 introns in the coding regions and only the DNA sequences in intron 4 are AT-rich. The close relationship between *Phaeosphaeria nodorum* and *Cochliobolus heterostrophus* was recently demonstrated in the encoded peptides of mating type genes (*MAT1-1* and *MAT1-2*) [2].

A fungal evolution study based on the phylogenetic analysis of several *gpd* gene nucleotide sequences was reported earlier [27]. Single-celled budding yeasts, such as *Saccharomyces cerevisiae* and *Zygosaccharomyces rouxii*, were hypothesized to have evolved separately from the filamentous ascomycetes, and both ascomycetes and basidiomycetes may have diverged much later than the budding yeasts and through a common ancestor [27]. In addition to coding sequences, differences in the number and size of introns in *gpd* genes may add data to the phylogenetic analyses. For example, the *gpd* gene in the ascomycete *Glomerella cingulata* having one intron and with 1,233bp in

length of coding region would be grouped with the basidiomycete *Cryptococcus curvatus* that contains the same length of the *gpd* gene coding region but contains 2 introns (Figure 3). However, these two fungi have only 72% peptide sequence similarity and are not closely related in the PAUP analysis (Figure 4). Even though *gpd* genes from many morphologically diversified fungi are not closely related at the nucleotide level, they grouped together phylogenetically based on amino acid sequence similarities (Figure 4). Phylogenetic analysis of nucleotide sequences from various organisms using different analytical software may construct a tree more suitable for explaining their potential evolution. Based on amino acid sequence conservation, it appears that most basidiomycetes evolve with budding yeasts, and many ascomycetes are associated with fission yeast and two basidiomycetes *Ustilago maydis* and *Coprinopsis cinerea*.

The strict conservation of GPD peptide sequences in three cereal *Phaeosphaeria* species indicates that they are closely related at this locus and that genetic divergence may have only occurred recently during their evolution. However, the nucleotide sequence diversity of the *gpd* genes supports previous attempts at species distinction based on RFLP fingerprints, rDNA ITS sequences and mating type gene conserved region sequences [7,29,30,31,32,33]. To study their evolution, the genetic diversity of other conserved genes in cereal *Phaeosphaeria* species will be further explored.

ACKNOWLEDGMENTS

We thank R.N.Cooley of Aventis Crop Science UK Ltd for providing the glyceraldehyde- 3-phosphate dehydrogenase gene sequence of *Phaeosphaeria nodorum* (BSm300 strain) to the GenBank (accession number AJ271155), Jia-shan Zhang of Wootton High School, Rockville, Maryland for her technical support, and B.M. Cunfer of University of Georgia and R.W. Hammond of USDA-ARS for their valuable suggestions.

LITERATURE CITED

1. Ballance, D. J. 1986. Sequences important for gene expression in filamentous fungi. Yeast 2:229-236.
2. Bennett, R. S., Yun, S. H., Lee, T. Y., Turgeon, B. G., Arseniuk, E., Cunfer, B. M., and Bergstrom, G. C. 2003. Identity and conservation of mating type genes in geographically diverse isolates of *Phaeosphaeria nodorum*. Fungal Genet. Biol. 40:25-37.
3. Berbee, M. L., Pirseyedi, M., and Hubbard, S. 1999. *Cochliobolus* phylogenetics and the origin of known, highly virulent pathogens, inferred from ITS and glyceraldehyde-3- phosphate dehydrogenase gene sequences. Mycologia 91:964-977.
4. Bitter, G. A., and Egan, K. M. 1984. Expression of heterologous genes in *Saccharomyces cerevisiae* from vectors utilizing the glyceraldehyde-3-phosphate dehydrogenase gene promoter. Gene 32:263-274.
5. Camara, M. P. S., O'Neill, N. R., and van Berkum, P. 2002. Phylogeny of *Stemphylium* spp. based on ITS and glyceraldehyde-3-phosphate dehydrogenase gene sequences. Mycologia 94:660-672.
6. Christiansen, S. K., Justesen, A. F., and Giese, H. 1997. Disparate sequence characteristics of the *Erysiphe graminis* f. sp. *hordei* glyceraldehyde-3-phosphate dehydrogenase gene. Curr. Genet. 31:525-529.
7. Cunfer, B. M., and Ueng, P. P. 1999. Taxonomy and identification of *Septoria* and *Stagonospora* species on small-grain cereals. Annu. Rev. Phytopathol. 37:267-284.
8. Denda, T., Kosuge, K., Watanabe, K., Ito, M., Suzuki, Y., Short, P. S., and Yahara, T. 1995. Intron length variation of the *Adh* gene in *Brachyscome* (Asteraceae). Plant Mol. Biol. 28:1067-1073.
9. Dusabenyagasani, M., Laflamme, G., and Hamelin, R. C. 2002. Nucleotide polymorphisms in three genes support host and geographic speciation in tree pathogens belonging to *Gremmeniella* spp. Can. J. Bot. 80:1151-1159.
10. Gilbert, W., Marchionni, M., and McKnight, G. 1986. On the antiquity of introns. Cell 46:151-154.
11. Harmsen, M. C., Schuren, F. H. J., Moukha, S. M., van Zuilen, C. M., Punt, P. J., and Wessels, J. G. H. 1992. Sequence analysis of the glyceraldehyde-3-phosphate dehydrogenase genes from the basidiomycetes *Schizophyllum commune*, *Phanerochaete chrososporium* and *Agaricus bisporus*. Curr. Genet. 22:447-454.
12. Holland, M. J., and Holland, J. P. 1978. Isolation and identification of yeast messenger ribonucleic acids coding for enolase, glyceraldehyde-3-phosphate dehydrogenase and phosphoglycerate kinase. Biochemistry 17:4900-4907.
13. Isoda, K., and Shiraishi, S. 2001. Allelic sequence polymorphisms in the intron region of the nuclear-encoded *GapC* gene preceeded the speciation of three closely related *Abies* species (Pinaceae). Theor. Appl. Genet. 102:244-250.
14. Jagadeeswaran, P., Forget, B. G., and Weissman, S. M. 1981. Short interspersed repetitive ENA elements in eucaryotes: transposable DNA elements generated by reverse transcription of RNA Pol III transcripts? Cell 26:141-142.
15. Jungehulsing, U., Arntz, C., Smit, R., and Tudzynski, P. 1994. The *Claviceps purpurea* glyceraldehyde-3-phosphate dehydrogenase gene: cloning, characterization, and use for the improvement of a dominant selection system. Curr. Genet. 25:101-106.
16. Kersanach, R., Brinkmann, H., Liaud, M., Zhang, D., Martin, W., and Cerff, R. 1994. Five identical intron positions in ancient duplicated genes of eubacterial origin. Nature 367:387-389.

17. Kinnaird, J. H., and Fincham, J. R. S. 1983. The complete nucleotide sequence of the *Neurospora* crassa *am* (NADP-specific glutamate dehydrogenase) gene. Gene 26:253- 260.
18. Krebs, E. G., Rafter, G. W., and Junge, J. M. 1953.Yeast glyceraldehyde-3-phosphate dehydrogenase. Yeast protein II. J. Biol. Chem. 200:479-492.
19. Kvarnheden, A., Tandre, K., and Engstrom, P. 1995. A *cdc2* homologue and closely related processed retropseudogenes from Norway spruce. Plant Mol. Biol. 27:391-403.
20. Langford, C. J., and Gallwitz, D. 1983. Evidence for an intron-contained sequence required for the splicing of yeast RNA polymerase II transcripts. Cell 33:519-527.
21. Michels, P. A. M., Marchand, M., Kohl, L., Allert S., Wierenga K, and Opperdoes, F.R. 1991. The cytosolic and glycosomal isoenzymes of glyceraldehyde-3-phosphate dehydrogenase in *Trypanosoma brucei* have a distant evolutionary relationship. Eur. J. Biochem. 198:421-428.
22. Osiewacz, H. D., and Ridder, R. 1991. Genome analysis of imperfect fungi: electrophoretic karyotyping and characterization of the nuclear gene coding for glyceraldehyde-3-phosphate dehydrogenase (*gpd*) of *Curvularia lunata*. Curr. Genet. 20:151-155.
23. Quigley, F., Martin, W. F., and Cerff, R. 1988. Intron conservation across the prokaryote-eukaryote boundary: structure of the nuclear gene for chloroplast glyceraldehyde-3-phosphate dehydrogenase from maize. Proc. Natl. Acad. Sci. USA 85:2672-2676.
24. Scharen, A. L., and Sanderson, F. R. 1985. Identification, distribution and nomenclature of the *Septoria* species that attack cereals. p.37-41 in "*Septoria* of cereals" by Scharen, A. L., (ed.), USDA-ARS, ARS-12, 116p.
25. Schubert, R., Sperisen, C., Muller-Starck, G., La Scala, S., Ernst, D., Sandermann, H., Jr., and Hager, K.P. 1998. The cinnamyl alcohol dehydrogenase gene structure in *Picea abies* (L.) Karst.: genomic sequences, Southern hybridization, genetic analysis and phylogenetic relationships. Trees 12:453-463.
26. Shah, D., Bergstrom, G. C., and Ueng, P. P. 1995. Initiation of Septoria nodorum blotch epidemics in winter wheat by seedborne *Stagonospora nodorum*. Phytopathology 85:452-457.
27. Smith, T. L. 1989. Disparate evolution of yeasts and filamentous fungi indicated by phylogenetic analysis of glyceraldehyde-3-phosphate dehydrogenase genes. Proc. Natl. Acad. Sci. USA 86:7063-7066.
28. Sprague, R. 1950. Diseases of cereals and grasses in North America. Ronald Press, New York.
29. Ueng, P. P., Bergstrom, G. C., Slay, R.M., Geiger, E. A., Shaner, G., Scharen, A.L. 1992. Restriction fragment length polymorphisms in the wheat glume blotch fungus, *Phaeosphaeria nodorum*. Phytopathology 82:1302-1305.
30. Ueng, P. P., and Chen, W. 1994. Genetic differentiation between *Phaeosphaeria nodorum* and *P. avenaria* using restriction fragment length polymorphisms. Phytopathology 84:800-806.
31. Ueng, P. P., Cunfer, B. M., Alano, A. S., Youmans, J. D., and Chen, W. 1995. Correlation between molecular and biological characters in identifying the wheat and barley biotypes of *Stagonospora nodorum*. Phytopathology 85:44-52.
32. Ueng, P. P., Dai, Q., Cui, K., Czembor, P. C., Cunfer, B. M., Tsang, H., Arseniuk, E., and Bergstrom, G. C. 2003. Sequence diversity of mating-type genes in *Phaeosphaeria avenaria*. Curr. Genet. 43:121-130.
33. Ueng, P. P., Subramaniam, K., Chen, W., Arseniuk, E., Wang, L., Cheung, A.M., Hoffmann, G. M., and Bergstrom, G. C. 1998. Intraspecific genetic variation of *Stagonospora avenae* and its differentiation from *S. nodorum*. Mycol. Res. 102:607-614.
34. Unkles, S. E., Logsdon, J. M., Jr., Robison, K., Kinghorn, J.R., and Duncan, J. M. 1997. Th *tigA* gene is a transcriptional fusion of glycolytic genes encoding triose-phosphate isomerase and glyceraldehyde-3-phosphate dehydrogenase in Oomycota. J. Bacteriol. 179:6816-6823.
35. Van den Hondel, C. A. M. J. J., and Punt, P. J. 1991. Gene transfer systems and vector development for filamentous fungi. In Applied Molecular Genetics of Fungi (J. F. Peberdy, C. E. Caten, J. E. Ogden and J. W. Bennett, eds.), p.1-28. Cambridge Univ. Press, Cambridge, UK. 187pp
36. Van Wert, S. L., and Yoder, O. C. 1992. Structure of the *Cochliobolus heterostrophus* glyceraldehyde-3-phosphate dehydrogenase gene. Curr. Genet. 22:29-35.
37. Waterham, H. R., Digan, M. E., Koutz, P. J., Lair, S. V., and Cregg, J. M. 1997. Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. Gene 186:37-44.

摘 要

翁溥[1,5]、E. Reszka[2]，鍾光仁[3], E. Arseniuk [2] and J. M. Krupinsky [4]. 2003. 穀類葉枯病菌 *Phaeosphaeria nodorum* 及 *P. avenaria* 之3-磷酸甘油醛脫氫酵素基因的比較. 植病會刊12:255-268. ([1.] 美國農業部 (USDA) 分子植物病理研究室；[2.] 波蘭植物育種及馴化研究所植物病理學系；[3.] 美國佛羅里達大學柑橘類作物研究及教育中心；[4.] 美國農業部北方大草原研究室；[5.] 聯絡作者：電子郵件 uengp@ba.ars.usda.gov；傳眞：+0021-301-504-5449）

穀類葉枯病 (Stagonospora leaf blotch of cereals) 是由 *Phaeosphaeria* 屬多種眞菌引起的複合病害。本研究利用 *P. nodorum*, *P. avenaria* f. sp. *triticea* (P.a.t.) 及 *P. avenaria* f. sp. *avenaria* (P.a.a.) 三種常見的穀類葉枯病菌的3-磷酸甘油醛脫氫酵素 (glyceraldehyde-3-phosphate dehydrogenase, GPD) 基因 (*gpd*) 探討它們的遺傳歧異度。利用 PCR 增幅上述三種穀類葉枯病菌的 *gpd* 基因，發現皆含有四個內含子 (intron)；在小麥 *P. nodorum* 分離株之 *gpd* 基因長度爲 1,253 至 1,255 個鹽基對 (bp) 不等。依據這些基因的第4 個內含子 (intron 4) 之序列差異，可將 *P. nodorum* 小麥分離株分爲二個亞群 (subgroups)；調查它們田間族群的分布，顯示這二個亞群間並無差異。除了 Pat2 分離株以外，P.a.t. 群之 Pat1 與 Pat3 分離株、P.a.a. 群菌株及大麥 *P. nodorum* 分離株的 *gpd* 基因長度皆爲 1,251 bp，且彼此演化遺傳分類關係頗爲相近。此外，*gpd* 基因的核酸變化大多發生在內含子區域及胺基酸密碼子的第3 個鹽基處，因此，除了 Pat2 分離株以外，上述三個穀類葉枯病菌的 gpd 基因，其推縯胺基酸序列均完全一致，至於 Pat2 分離株的 GPD 蛋白質則含有 2 個胺基酸的差異。研究顯示前述三種穀類葉枯病菌的 GPD 蛋白序列和玉米病原菌 *Cochliobolus heterostrophus* 的 GPD 蛋白序列相當接近。根據演化遺傳分類關係之分析結果，本文中也將進一步討論這些穀類葉枯病菌與其他子囊菌間的演化關係。

關鍵詞：3-磷酸甘油醛脫氫酵素基因、*Phaeosphaeria* 屬、穀類葉枯病、小麥