Mycoscience 40: 517–523, 1999

Cloning and sequence analysis of the glyceraldehyde-3-phosphate dehydrogenase gene from the ectomycorrhizal basidiomycete *Lyophyllum shimeji*

Takeshi Saito and Norio Tanaka

The Mushroom Research Institute of Japan, 8–1 Hirai-cho, Kiryu, Gunma 376–0051, Japan

Accepted for publication 24 September 1999

The *GPD* gene encoding glyceraldehyde-3-phosphate dehydrogenase was isolated and characterized from the ectomycorrhizal basidiomycete *Lyophyllum shimeji*. This gene was a single copy and had a coding capacity of 337 amino acids interrupted by nine introns. The deduced amino-acid sequence of the protein encoded by this gene was highly homologous to those of the *GPD* genes from the saprophytic fungi *Schizophyllum commune*, *Phanerochaete chrysosporium*, and *Agaricus bisporus*, and the ectomycorrhizal fungi *Amanita muscaria*, *Boletus edulis*, and *Lactarius deterrimus*. The promoter region of *L. shimeji GPD* gene contained a CCAAT box, a TATAAAA box, and a CT-stretch. The major transcriptional initiation site was located 31 nucleotides downstream from the TATAAAA box and in the CT-stretch.

Key Words——cloning; glyceraldehyde-3-phosphate dehydrogenase; *Lyophyllum shimeji*.

Glyceraldehyde-3-phosphate dehydrogenase (GPD or GAPDH, E.C.1.2.1.12) is a key enzyme in glycolysis. In expectation of high and constitutive expression, the genes of this enzyme were cloned from *Schizophyllum commune* Fr.: Fr., *Phanerochaete chrysosporium* Burds, and *Agaricus bisporus* (J. Lange) Imbach var. *albidus* (J. Lange) Sing. to get a strong promoter of transformation vector for basidiomycetes (Harmsen et al., 1992). Besides these saprophytic fungi, PCR amplification of *GPD* genes was tried for ectomycorrhizal basidiomycetes, *Amanita muscaria* (L.: Fr.) Hooker, *Boletus edulis* Bull.: Fr., and *Lactarius deterrimus* Gröger (Kreuzinger et al., 1996), but the amplified DNA products had partial sequences for the genes.

Though *Lyophyllum shimeji* (Kawamura) Hongo is an ectomycorrhizal fungus, Ohta (1994) reported mature fruit-body formation with an artificial medium and without a host plant. This reveals that *L. shimeji* has both symbiotic and saprophytic characters. From these points of view, it is expected that the transcriptional control regions (promoter and terminator) of the *L. shimeji GPD* gene can work in cells of both ectomycorrhizal and saprophytic basidiomycetes. The gene was cloned and sequenced with a view to constructing an excellent transformation vector.

Materials and methods

PCR amplification of the *GPD* gene and subcloning *Lyophyllum shimeji* MRI 6532 was isolated from a wild mushroom in Gunma prefecture and used throughout this work. The organism was cultured and maintained in GMY liquid medium (1% (w/v) glucose, 1% (w/v) malt extract, 0.4% (w/v) yeast extract). Genomic DNA was prepared from freeze-dried mycelium by extraction and purification using the method of Murray and Thompson (1980) with minor modification.

Primer-U (5′-GG**AAGCTT**GGTCTACATGTTCAAGTACGAC-3′: AAGCTT is the *Hin*dIII linker) and primer-D (5′-GG**GGTACC**TAGCCCCACTCGTTGTCGTCC-3′: GGTACC is the *Kpn*I linker), whose corresponding amino-acid sequences were VYMFKYD and WYDNEWGY, respectively, were designed to amplify a part of the *GPD* gene. The reaction mixture for PCR amplification was prepared according to the guide of *TaKaRa Ex Taq*™ (Takara Shuzo, Osaka, Japan). The thermal cycle controller, Program Temp. Control System PC-700 (Astec, Fukuoka, Japan), was set as follows: a single cycle of denaturation for 5 min at 94°C followed by 35 cycles of 1 min at 94°C, 1 min at 55°C, and 1 min at 72°C, and finally a single cycle of 5 min at 72°C. The holding temperature was 4°C.

The amplified DNA fragment was digested with two restriction enzymes, *Hin*dIII and *Kpn*I. The digested fragment was inserted into dephosphorylated *Hin*dIII- and *Kpn*I-digested pUC18 plasmid vector, followed by transformation of *E. coli* JM109 (Yanisch-Perron et al., 1985). The subcloned fragments were sequenced by a Circum Vent™ Thermal Cycle Sequencing Kit (New England Biolabs, Beverly, USA).

Southern blot analysis to detect the complete *GPD* gene *Lyophyllum shimeji* genomic DNAs were digested with one of several restriction enzymes, *Hin*dIII, *Eco*RI, *Pst*I, *Xba*I, or *Bam*HI, and electrophoresed in 1% (w/v) agarose gel (Agarose L 03: Takara Shuzo, Osaka, Japan), then transferred onto nylon membrane (Hybond™-N: Amer-

sham Pharmacia Biotech UK, Bucks., UK).

Hybridization was done with a biotin-labeled probe (prepared by NEBlot Phototope Kit: New England Biolabs, Beverly, USA) of the subcloned *GPD* partial sequence in a solution of 6×SSC (1×SSC: 0.15 M NaCl, 0.015 M Na_3-citrate, pH 7.0), 0.5% (w/v) SDS, 10×Denhardt's solution (50×Denhardt's solution: 1% (w/v) Ficoll, 1% (w/v) polyvinylpyrolidone, 1% (w/v) bovine serum albumin), and 0.2 mg/ml denatured salmon sperm DNA at 68°C, overnight. The membrane was washed with 2×SSC twice at room temperature, 0.1×SSC containing 0.1% (w/v) SDS at 68°C for 1 h, and 2×SSC once at room temperature. The hybridization signals were detected by chemiluminescence with a Phototope™-Star Detection Kit (New England Biolabs, Beverly, USA).

Isolation of the *GPD* gene from genomic DNA libraries *Lyophyllum shimeji* genomic DNA was digested with *Hin*dIII, and DNA fragments of 3 to 4 kb and 1.5 to 2 kb were recovered. The two sets of purified genomic DNA fragments were inserted into dephosphorylated *Hin*dIII-digested pUC18 plasmid vectors, followed by transformation of *E. coli* JM109 to construct two genomic DNA libraries.

The libraries were screened for the *GPD* gene sequences by colony hybridization (Grunstein and Hogness, 1975) with a biotin-labeled probe of the subcloned *GPD* partial sequence. Hybridization and washing of membranes were done by the same method as in the Southern blot analysis.

Restriction map, sequencing and primer extension analysis The positive clones isolated from the two libraries were digested with one of several restriction enzymes, *Hin*dIII, *Bgl*II, *Nco*I, *Pst*I, *Xho*I, or *Hin*cII. The digestion patterns and the restriction sites of each clone were examined by 1% (w/v) agarose gel (Agarose L 03) electrophoresis. The nucleotide sequence of each cloned DNA fragment was determined by use of a Circum Vent™ Thermal Cycle Sequencing Kit. Promoter and terminator regions were confirmed by Southern blot analysis of the digest and sequencing of each cloned DNA fragment.

A transformation vector was constructed with a *Hin*cII-*Nco*I fragment containing the promoter region of the *GPD* gene (see Fig. 1), the hygromycin B phosphotransferase gene (*hph* gene) of *E. coli* (Gritz and Davies, 1983), and an *Xho*I-*Hin*cII fragment containing the terminator region of the *GPD* gene (see Fig. 1). Protoplasts of *Lentinula edodes* (Berk.) Pegler were transformed by electroporation using Gene Pulser™ (Bio-Rad laboratories, Richmond, CA, USA), and transformants were selected on 0.6 M SMY agar plates (0.6 M sucrose, 1% (w/v) malt extract, 0.4% (w/v) yeast extract, 1.5% (w/v) agar) containing 50 μg/ml hygromycin B (Wako Pure Chemical Industries, Osaka, Japan). The details of the transformation techniques will be reported elsewhere.

Total RNA was extracted from mycelium of one transformant by the guanidinium-thiocyanate method (Maniatis et al., 1989), and mRNA was purified from the total RNA with Oligotex-dT 30 <Super> (Takara Shuzo, Osaka, Japan). The poly(A)$^+$ RNA (4 μg) was incubated with 50 pmol of a synthetic 27 mer biotin-labeled oligonucleotide (5'-biotin-CTCGACACACGTCGCGGTGAGTTCAGG-3') complementary to the inner sequence of the *hph* gene at 37°C for 15 min in reverse transcription buffer (50 mM Tris-HCl (pH 8.0), 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 0.5 mM each deoxyribonucleotide triphosphate). Then 200 units of Super Script™ reverse transcriptase (Gibco BRL, Life Technologies, Rockville, MD, USA) and 0.4 mM biotin-14-dATP (Amersham Pharmacia Biotech UK, Bucks., UK) were added to the mixture and further incubated at 37°C for 45 min. To determine the transcriptional initiation sites, the 5'-flanking region of the *hph* gene (promoter region of *L. shimeji GPD* gene) was sequenced with the same biotin-labeled oligonucleotide as the sequencing primer. The primer extension reaction mixture and sequence reaction mixture were electrophoresed together in 6% (w/v) polyacrylamide gel with 7 M urea. The primer extension products and the sequence ladders were detected by chemiluminescence with a Phototope™-Star Detection Kit.

Computer analyses The nucleotide and derived amino-acid sequences were analyzed with the programs for gene analysis, GENETYX-WIN Ver. 3 (Software Development, Tokyo, Japan). The amino-acid sequence and the positions of introns were determined by comparison with some known GPD amino-acid sequences and intron consensus sequences (Harmsen et al., 1992 and Kreuzinger et al., 1996). Multiple sequence alignment was done with multiple alignment programs (Barton and Sternbert, 1987) of GENETYX-WIN Ver. 3. Some sequence data of the *GPD* genes were obtained from the database of DDBJ.

Nucleotide sequence accession numbers The nucleotide sequence data of the *L. shimeji GPD* gene and its 5'-flanking region reported in this paper have been assigned DDBJ accession numbers D88426 and AB028458, respectively.

Results

Restriction map and sequence analysis of *L. shimeji GPD* gene Two signals corresponding to DNA fragments of about 3.7 and 1.7 kb were detected by the Southern blot analysis for *Hin*dIII-digested genomic DNA using a biotin-labeled probe of the subcloned *GPD* partial sequence. To clone them, two genomic DNA libraries containing DNA fragments of 3 to 4 kb and 1.5 to 2 kb were constructed and screened for the *GPD* gene sequences by

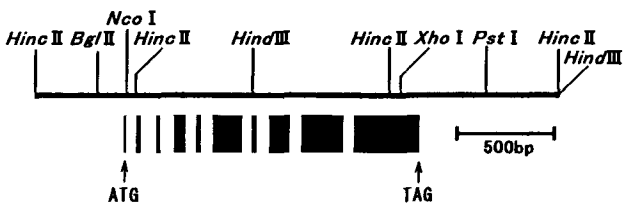

Fig. 1. Restriction map of the DNA fragment containing the coding regions and the transcriptional control regions of *Lyophyllum shimeji GPD* gene.
Exons are indicated under the map by black boxes and gaps between exons are introns.

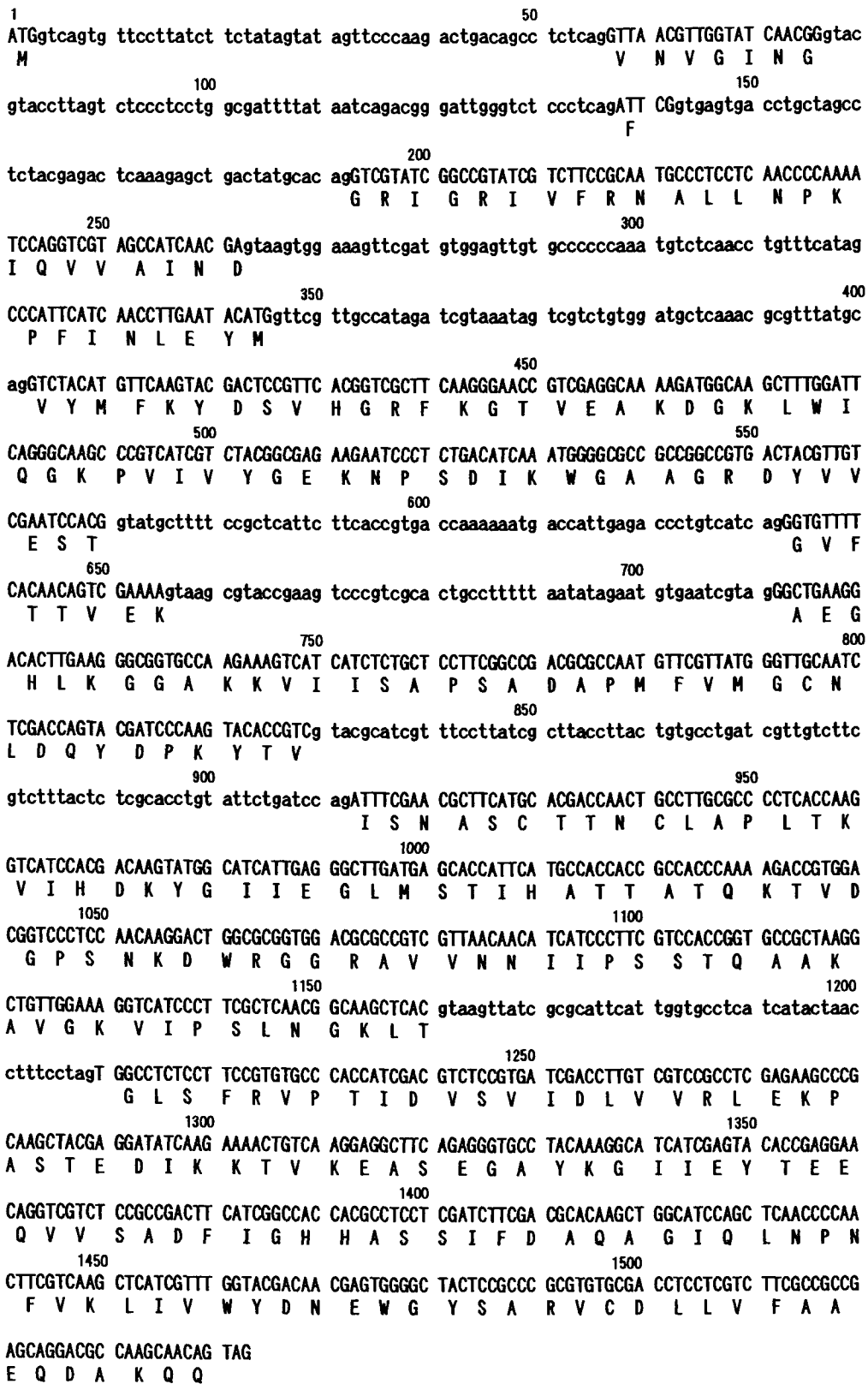

Fig. 2. Nucleotide sequence and deduced amino-acid sequence of *Lyophyllum shimeji GPD* gene.
Exon sequences are capitalized and intron sequences are indicated by small letters. Size of the gene was 1543 bp and the putative protein consisted of 337 amino acids.

the colony hybridization method. Two and four positive clones were isolated from the two libraries, respectively. Examination of the digestion pattern and restriction sites of each clone showed that both of the former clones had the same insertions of about 3.0 kb DNA fragments and all of the latter clones contained the same insertions of about 1.7 kb DNA fragments. The Southern blot analysis and sequencing revealed that the 3.0 kb DNA fragment contained the promoter region and the 5′-terminal sequence of the *GPD* gene with the deduced translational start codon, ATG. On the other hand, the 1.7 kb DNA fragment had the terminator region and the 3′-terminal sequence of the *GPD* gene with the deduced translational stop codon, TAG.

The restriction map is shown in Fig. 1. The nucleotide sequence of the *GPD* structure gene was 1543 bp, and the coding region of a putative protein consisting of 337 amino acids was interrupted by nine introns (Fig. 2).

Nucleotide sequence of the promoter region and determination of the transcriptional initiation sites The sequence of 418 nucleotides preceding the predicted translational start codon is shown in Fig. 3a. This region contained a TATAAAA box, a CCAAT box (Breathnach and Chambon, 1981), and a CT-rich stretch (Gurr et al., 1988). A transformation vector using transcriptional control regions (promoter and terminator) of the *L. shimeji*

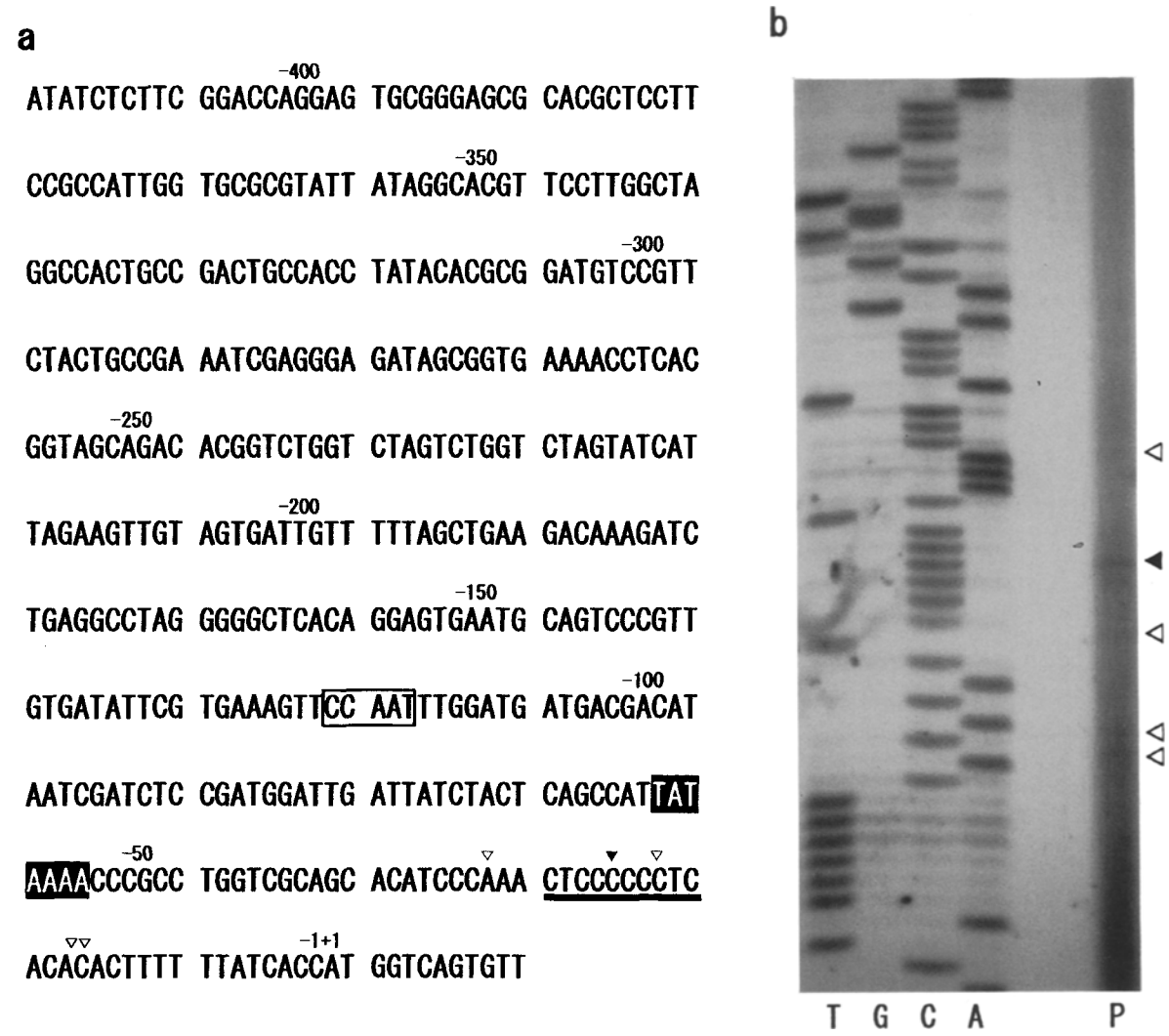

Fig. 3. Structure of the promoter region and determination of the transcriptional initiation sites of *Lyophyllum shimeji GPD* gene.
a. The nucleotide sequence (418 bp) of the 5′-flanking region of the *GPD* structure gene. TATAAAA box is indicated by inverted letters and CCAAT box is surrounded. CT-stretch is underlined. One major and four alternative transcriptional initiation sites, which were determined by primer extension method, are indicated by filled and open triangles, respectively. The numbers above the nucleotide sequence indicate the positions of nucleotides relative to the start codon.
b. Determination of the transcriptional initiation sites of the *GPD* gene by primer extension method. The stop positions of the primer extension reaction were determined from the sequencing ladders (indicated by T, G, C, A), which were derived from a transformation vector with the 5′-flanking region of the *GPD* gene and *hph* gene, by dideoxy chain-termination method using a 27 base primer as the sequencing primer. One filled and four open triangles indicate the primer extension products (indicated by P) corresponding to major and alternative transcriptional initiation sites, respectively.

GPD gene and the *hph* gene of *E. coli*, which has no introns, was constructed in order to analyze the transcriptional initiation sites by the primer extension method (McKnight and Kingsbury, 1982). *Lentinula edodes* was transformed with this vector (data not shown), and primer extension analysis was done with poly(A)$^+$ RNA purified from mycelium of a transformant. As shown in Fig. 3b, one strong and four weak bands were observed at the positions corresponding to nucleotide −24 (C-residue) and nucleotides −31 (A-residue), −21 (C-residue), −16 (A-residue), and −15 (C-residue), respectively. The major transcriptional initiation site was located 31 nucleotides downstream from the TATAAAA box, just in the CT-rich sequence (see Fig. 3a).

Amino-acid sequence of *L. shimeji* GPD protein Aligned amino-acid sequences of the GPD proteins from the nucleotide sequence data are shown in Fig. 4. The *L. shimeji* GPD protein was compared with those of three saprophytic fungi (*S. commune*, *P. chrysosporium* and *A. bisporus*) and three ectomycorrhizal fungi (*A. muscaria*, *L. deterrimus* and *B. edulis*). The amino-acid sequence data of saprophytic fungi were complete, but those of ectomycorrhizal fungi were partial. The amino-acid sequence of *L. shimeji* showed about 74 to 76% homology with those of the three saprophytic fungi, other than *A. bisporus*-1, which could be the product of a pseudogene (Harmsen et al., 1992). The *L. shimeji* GPD protein was also about 74 to 77% homologous to those of three ectomycorrhizal fungi. The amino-acid sequences of the GPD protains were highly conserved among these fungi.

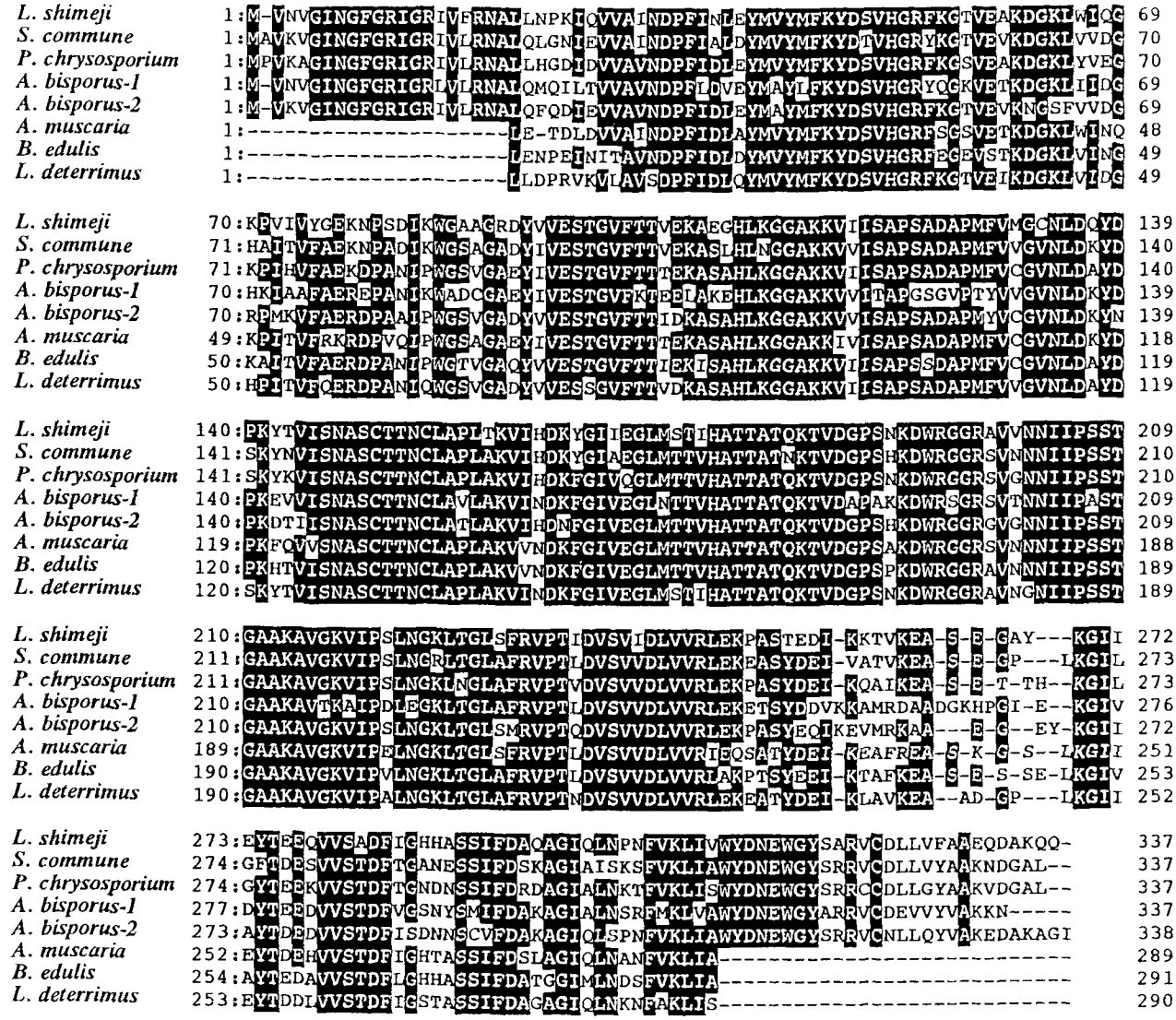

Fig. 4. Aligned amino-acid sequences of the GPD proteins from *Lyophyllum shimeji*, *Schizophyllum commune*, *Phanerochaete chrysosporium*, *Agaricus bisporus* (1 and 2), *Amanita muscaria*, *Boletus edulis*, and *Lactarius deterrimus*.
Identical residues in more than four of eight sequences are represented in inverted letters. The sequences from *A. muscaria*, *B. edulis*, and *L. deterrimus* are partial.

Discussion

Structure of the promoter region The promoter is the most important region in the construction of a transformation vector. In expectation of high and constitutive expression, the promoters of *GPD* genes have been used in several transformation vectors of fungi, e.g., *Aspergillus nidulans* (Eidam) Winter (Punt et al., 1987), *Podospora anserina* (Ces.) Niessl (Ridder et al., 1992), *S. commune* (Schuren et al., 1993; Schuren and Wessels, 1994), and *A. bisporus* (Rhee et al., 1996). *S. commune GPD* gene has a consensus TATA box, TATAAAA, but does not have a CCAAT box in its promoter. *GPD* genes of *P. chrysosporium*, *A. bisporus*-1, and *A. bisporus*-2 (*A. bisporus* has two *GPD* genes) contain untypical TATA boxes, TTAAAT, TACAAA and TACAAAAA, respectively. All of them have a CCAAT or CAAT box (Harmsen et al., 1992). The promoter of *L. shimeji GPD* gene contains a typical TATA box (TATAAAA) and a CCAAT box. Though *GPD* promoters have various structures in these fungi, major transcriptional initiation sites are located the estimated positions of 24 to 32 nucleotides downstream from TATA box.

CT-stretches are located downstream from the TATA boxes in all the promoter regions. *Schizophyllum commune* and *P. chrysosporium* have two CT-stretches, but *A. bisporus* and *L. shimeji* have only one, and that of *L. shimeji* is especially short (10 bp: see Fig. 3a). It remains to be studied how the number and length of CT-stretches influence the promoter activity, especially the strength of expession.

Exon-intron patterns As all introns are of almost constant size (about 50 to 80 bp), the appearances of *GPD* structure genes can be represented by exon positions and sizes (exon-intron patterns: Fig. 5). In the figure, *L. shimeji*, *S. commune*, *P. chrysosporium*, *A. bisporus* (1 and 2), *A. muscaria*, *B. edulis*, and *L. deterrimus* are compared. Nucleotide sequence data of *L. shimeji* and saprophytic fungi (*S. commune*, *P. chrysosporium* and *A. bisporus*) are complete, but those of ectomycorrhizal fungi (*A. muscaria*, *B. edulis* and *L. deterrimus*) are partial. The pattern of *L. shimeji GPD* gene is quite similar to those of *A. bisporus*-1 and *A. bisporus*-2.

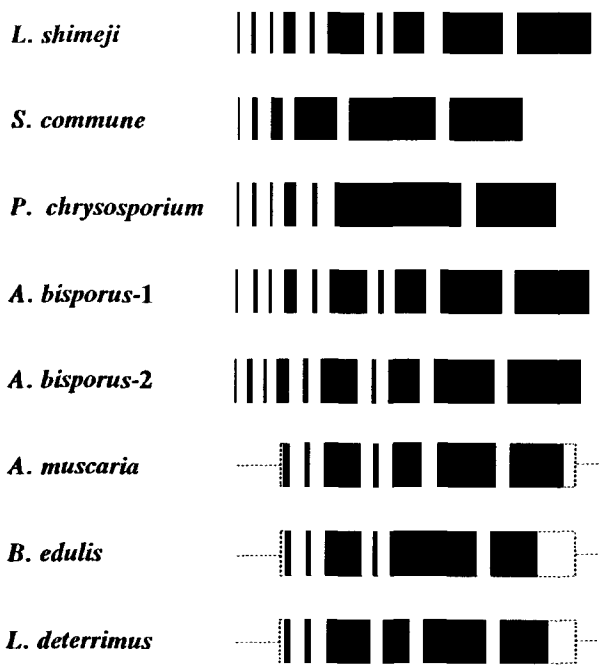

Fig. 5. Comparison of exon-intron patterns in *GPD* genes. Black boxes represent the positions and sizes of exons and gaps between the black boxes indicate introns.

Kreuzinger et al. (1996) reported that the intron pattern of *L. deterrimus* was similar to that of *A. bisporus*, but they are appreciably different in our exon-intron patterns. *Amanita muscaria* is more closely related to *A. bisporus* than is *L. deterrimus* in the exon-intron patterns. The relationship of fungi is sometimes discussed by comparison of intron positions (intron patterns) in *GPD* genes, and it has been suggested that the intron pattern might be a potential taxonomic marker (Harmsen et al., 1992; Kreuzinger et al., 1996). We use the exon-intron pattern indicated by exon positions and sizes, since it gives more accurate data than intron pattern as a taxonomic marker.

The complete *GPD* gene of *L. shimeji* was cloned and its sequence was analyzed. The gene was 1543 bp and encoded a protein consisting of 337 amino acids, whose amino-acid sequence was highly homologous to those of several basidiomycetes. To determine the transcriptional initiation sites by the primer extension method, we constructed a transformation vector with transcriptional control regions of the *GPD* gene. This vector might be expected to work well in cells of both ectomycorrhizal and saprophytic basidiomycetes. We are planning to try it in transformations of several mushrooms and their applications.

Acknowledgements——This investigation was supported by Gunma Prefectural Forestry Research Laboratory and Mori & Company, Ltd. We gratefully acknowledge Dr. G. Kawai for significant advice.

Literature cited

Barton, J. G. and Sternbert, M. J. 1987. A strategy for the rapid multiple alignment of protein sequences: confidence levels from tertiary structure comparisons. J. Mol. Biol. **198**: 327–337.

Breathnach, R. and Chambon, P. 1981. Organization and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. **50**: 349–383.

Gritz, L. and Davies, J. 1983. Plasmid encoding hygromycin B resistance: the sequence of hygromycin B phosphotransferase gene and its expression in *Escherichia coli* and *Saccharomyces cerevisiae*. Gene **25**: 179–188.

Grunstein, M. and Hogness, D. S. 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. USA **72**: 3961–3965.

Gurr, S. J., Unkles, S. E. and Kinghorn, J. R. 1988. The structure and organization of nuclear genes of filamentous fungi. In: "Gene Structure in Eukaryotic Microbes.," (ed. by Kinghorn J. R.), SGM Special Publication, vol. 23, pp. 93–139. IRL Press, Oxford.

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 chrysosporium* and *Agaricus bisporus*. Curr. Genet. **22**: 447–454.

Kreuzinger, N., Podeu, R., Gruber, F., Göbl, P. and Kubicek, C. P. 1996. Identification of some ectomycorrhizal basidiomycetes by PCR amplification of their *gpd* (glyceraldehyde-3-phosphate dehydrogenase) genes. Appl. Environ. Microbiol. **62**: 3432–3438.

Maniatis, T., Fritsch, E. F. and Sambrook, J. 1989. "Molecular Cloning", 2nd ed., pp. 7.19–7.22. Cold Spring Harbor Laboratory Press, New York.

McKnight, S. L. and Kingsbury, R. 1982. Transcriptional control signals of a eukaryotic protein-coding gene. Science **217**: 316–324.

Murray, M. G. and Thompson, W. F. 1980. Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. **8**: 4321–4325.

Ohta, A. 1994. Production of fruit-bodies of a mycorrhizal fungus, *Lyophyllum shimeji*, in pure culture. Mycoscience **35**: 147–151.

Punt, P. J., Oliver, R. P., Dingemanse, M. A., Pouwels, P. H. and van den Hondel, C. A. M. J. J. 1987. Transformation of *Aspergillus* based on the hygromycin B resistance marker from *Escherichia coli*. Gene **56**: 117–124.

Rhee, M. D., Graça, P. M. A., Huizing, H. J. and Mooibroek, H. 1996. Transformation of the cultivated mushroom, *Agaricus bisporus*, to hygromycin B resistance. Mol. Gen. Genet. **250**: 252–258.

Ridder, R. and Osiewacz, H. D. 1992. Sequence analysis of the gene coding for glyceraldehyde-3-phosphate dehydrogenase (*gpd*) of *Podospora anserina*: use of homologous regulatory sequences to improve transformation efficiency. Curr. Genet. **21**: 207–213.

Schuren, F. H. J., Harmsen, M. C. and Wessels, J. G. H. 1993. A homologous gene-reporter system for the basidiomycete *Schizophyllum commune* based on internally deleted homologous genes. Mol. Gen. Genet. **238**: 91–96.

Schuren, F. H. J. and Wessels, J. G. H. 1994. Highly-efficient transformation of the homobasidiomycete *Schizophyllum commune* to phleomycin resistance. Curr. Genet. **26**: 179–183.

Yanisch-Perron, C., Vieira, J. and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**: 103–119.