

Molecular characterization and isolation of cytochrome P450 genes from the filamentous fungus *Aspergillus oryzae*

K. H. M. Nazmul Hussain Nazir · Hirofumi Ichinose · Hiroyuki Wariishi

Received: 11 December 2009 / Revised: 22 February 2010 / Accepted: 1 March 2010 / Published online: 1 April 2010
© Springer-Verlag 2010

Abstract We explored the molecular diversity of cytochrome P450 genes in the filamentous fungus *Aspergillus oryzae* using bioinformatic and experimental approaches. Based on bioinformatic annotation, we found 155 putative genes of cytochromes P450 in the whole genome sequence; however, 13 of 155 appeared to be pseudogenes due to sequence deletions and/or inframe stop codon(s). There are 87 families of *A. oryzae* cytochromes P450 (AoCYPs), indicating considerable phylogenetic diversity. To characterize *A. oryzae* AoCYPs, we attempted to isolate cDNAs using RT-PCR and determined their transcriptional capabilities. To date, we have confirmed gene expression of 133 AoCYPs and cloned 121 AoCYPs as full-length cDNAs encoding a mature open reading frame. Using experimen-

tally deduced sequences and intron–exon organization, we analyzed AoCYPs phylogenetically. We also identified intronic consensus sequences in AoCYPs genes. The experimentally validated exonic and intronic sequences will be a powerful advantage in identification and characterization of novel P450s from various ascomycetous fungi.

Keywords Cytochrome P450 · *Aspergillus oryzae* · Molecular diversity · cDNA cloning

Introduction

Cytochromes P450 (P450s) constitute a large superfamily of heme-containing monooxygenases, which are distributed in a wide variety of organisms. The large-scale molecular variations among P450 species imply an evolutionary trajectory in which a common ancestor extensively branched into various organisms (Gotoh 1993; Nelson et al. 1993; Lewis et al. 1998). The vast majority of P450s are thought to have specifically emerged and individually diversified during evolution of each organism. For instance, only CYP51, which is involved in the sterol biosynthesis pathway, is conserved across eukaryotic phyla (Yoshida 1993; Aoyama et al. 1996). Thus, the molecular and functional diversities of P450s enable specialization to meet the metabolic requirements of each organism, especially secondary metabolic pathways such as detoxification of xenobiotics and synthesis of secondary metabolites (Ortiz-de-Montellano 2005; Demain and Fang 2000). However, the degree of P450 divergence, such as the numbers of genes and gene families, differs significantly across biological kingdoms, phyla, and species. Such divergences presumably reflect an evolutionary driving force to develop survival strategies (Nelson 1999). In particular, the scale of divergence of

Communicated by Axel Brakhage.

Electronic supplementary material The online version of this article (doi:10.1007/s00203-010-0562-z) contains supplementary material, which is available to authorized users.

K. H. M. Nazmul Hussain Nazir ·
H. Ichinose (✉) · H. Wariishi (✉)
Faculty of Agriculture, Kyushu University,
6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan
e-mail: ichinose@agr.kyushu-u.ac.jp

H. Wariishi
e-mail: hirowari@agr.kyushu-u.ac.jp

H. Wariishi
Bio-Architecture Center, Kyushu University,
6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan

H. Wariishi
Innovation Center for Medical Redox Navigation,
Kyushu University, 6-10-1 Hakozaki,
Higashi-ku, Fukuoka 812-8581, Japan

P450s within the fungal kingdom is unprecedented (Park et al. 2008; Deng et al. 2007; Doddapaneni et al. 2005; Intikhab et al. 2007). In spite of their high sequence diversity, P450s share conserved overall protein architecture and contain several conserved sequences, such as the FxxGxxx-CxG signature motif. Such motifs can be used to discover novel genes from genomic databases. Within the last few years, the sequence database of P450s has enlarged exponentially, and continues to increase (<http://drnelson.utmem.edu/CytochromeP450.html>; <http://p450.riceblast.snu.ac.kr/index.php?a=view>; Park et al. 2008). A compilation of P450 sequences will increase understanding of metabolic diversity and evolutionary history of living organisms. However, there are few detailed studies on transcriptional profiles and catalytic functions of P450s in metabolic processes, so their exact roles and functions are still poorly understood.

A wide variety of filamentous fungi is used to produce economically valuable consumer items. The filamentous fungus *Aspergillus oryzae* is one of the most widely used microorganisms, and has been used for more than 1,000 years in Japanese fermentation industries to produce indigenous products such as *sake* (rice wine), *miso* (soybean paste), and *shoyu* (soy sauce). Because of the long history of use in fermentation and food production, *A. oryzae* is listed as “generally recognized as safe” by the food and drug administration in the United States. Besides fermentation technologies, many studies on production of recombinant enzymes and primary and secondary metabolites have focused on *A. oryzae* (Tailor and Richardson 1979; Abe et al. 2006). Recently, the whole genomic sequence of *A. oryzae* (strain RIB40) was determined and made available to the public (Machida et al. 2005). *A. oryzae* has eight chromosomes with a total genome size of 37.6 Mb, which is 20–30% larger than the genomes of *A. nidulans* and *A. fumigatus*. In total, there are 12,074 predicted genes in *A. oryzae*, compared with 9,396 in *A. nidulans* and 9,009 in *A. fumigatus* (Machida et al. 2005; Galagan et al. 2005; <http://www.bio.nite.go.jp/dogan/Top>). The increased gene number in *A. oryzae* is mainly due to a gain of extra genes involved in secondary metabolic pathways, including P450s, suggesting that *A. oryzae* has unique metabolic processes that are absent in other *Aspergillus* species. Thus, genomic data will increase molecular understanding of previously uncharacterized metabolic processes in *A. oryzae*.

In the present study, we explored the molecular diversity of *A. oryzae* P450s (AoCYPs) using a bioinformatic annotation and experimental validation. To our knowledge, this is the first comprehensive transcriptional survey of ascomycetous P450s. The identified and isolated AoCYPs have potential benefits to improve bioinformatic algorithms, expand biochemical knowledge, and advance biotechnology.

Materials and methods

Microorganism and culture conditions

Aspergillus oryzae strain RIB40 (NBRC 100959) was cultured on YPD/agar plates (1% yeast-extract, 2% bacto-peptone, 0.04% adenine sulfate, 2% glucose, 1.5% bacto-agar) at 30°C for 3 days, and then inoculated into synthetic culture media as described previously (Kirk et al. 1978). The synthetic media contained 1% glucose, 1.2 mM (nitrogen-limited conditions) or 12 mM (nitrogen-enriched conditions) ammonium tartrate, 20 mM dimethylsuccinate (pH 4.5), and trace elements (100 mL/L culture). The trace elements (1 L) contained 20 g KH_2PO_4 , 5.3 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 1.3 g $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 10 or 0 mg thiamine hydrochloride, 150 mg $\text{N}(\text{CH}_2\text{COOH})_2$, 18 mg $\text{CoSO}_4 \cdot 7\text{H}_2\text{O}$, 18 mg $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 1 mg $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 1.8 mg $\text{AlK}(\text{SO}_4)_{12}\text{H}_2\text{O}$, 1 mg H_3BO_3 , 70 mg $\text{MnSO}_4 \cdot 5\text{H}_2\text{O}$, 10 mg $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 100 mg NaCl, and 1 mg $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$. Fungal cells were grown with shaking (130 rpm) at 30°C for 2–21 days under aerobic conditions.

RNA extraction and first-strand cDNA synthesis

Total RNA was extracted individually from 5-, 10-, 18-, and 21-day-old mycelia using the acid guanidinium-phenol-chloroform method and further purified using an RNeasy Plant Mini Kit (QIAGEN). The concentration of RNA was calculated from the absorbance at 260 nm. Equal quantities of RNA isolated from mycelia of the four different ages were then mixed. The RNA cocktail was treated with DNase I (Takara), and first-strand cDNAs were synthesized with SUPERScript III™ reverse transcriptase (Invitrogen) in the presence of oligo(dT) primer (5'-TTTT TTTTTTTTTTTTTT-3'; V = A, C, or G). The reaction mixtures (50 μL) contained 5 μg total RNA, 200 units SUPERScript™ III reverse transcriptase (Invitrogen), 40 units RNase-Out (Invitrogen), 4 mM DDT, 0.4 mM dNTPs, and 25 pmol oligo(dT) primer in 1 \times first strand buffer, and were incubated at 50°C (60 min) for the extension reaction. The reaction mixtures were stored at -20°C until PCR amplifications.

Bioinformatic annotation of P450 from *A. oryzae*

A possible coding sequence for AoCYPs was used to search the National Institute of Technology and Evaluation database based upon sequence similarity to known P450s (<http://www.bio.nite.go.jp/dogan/Top>). To evaluate annotation accuracy, we identified the P450 signature sequence (F-x-x-G-x-x-x-C-x-G) in the heme-binding domain, the E-x-x-R motif in the K-helix, a conserved Thr in the center of the I-helix, and the hydrophobic transmembrane domain

(TMD) at the N-terminal region. TMD sequences were analyzed by the SOSUI (<http://bp.nuap.nagoya-u.ac.jp/sosui/>) and the TMHMM server v.2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>). If candidates lacked sequences corresponding to these regions, their capability to encode P450 was judged by overall sequence similarity to known P450s.

cDNA amplification by RT-PCR

PCR amplification were carried out using gene-specific primers designed to anneal to 5'- and 3'-untranslated regions; basically, to the 2–30 bp upstream or downstream flanking sequence from the putative start and stop codons (Supplemental Table). Custom-synthesized oligonucleotide primers were obtained from SIGMA-ALDRICH. cDNA was amplified by Phusion DNA polymerase (New England Biolabs). The reaction mixture (50 μ L) contained first-strand cDNA solution (1 μ L), dNTP (200 μ M), primers (2 μ M each), DMSO (2%), and Phusion DNA polymerase (0.02 U/ μ L) in Phusion HF buffer. The reaction conditions were programmed as follows: denaturation at 96°C for 3 min; 40 cycles of 96°C for 30 s, 55°C for 20 s, and 72°C for 60 s; and final extension at 72°C for 2 min. The PCR products were separated by 1.5% agarose gel electrophoresis and visualized on a UV-transilluminator. Target cDNAs were purified using a QIAquick Gel Extraction Kit (QIAGEN), phosphorylated with T4 polynucleotide kinase (TaKaRa), cloned into the pUC18 *Sma*I site, and transformed into *Escherichia coli* strain JM109. Positive transformants were selected on LB-agar plates containing ampicillin (100 mg/L), IPTG (100 mg/L), and X-gal (100 mg/L). The positive clones were further grown in LB medium supplemented with ampicillin (100 mg/L). Plasmids harboring AoCYP cDNA were extracted using a QIAprep Spin Miniprep Kit (QIAGEN) and sequenced with an automated DNA Sequencer (CEQ 8000; Beckman) using a DTCS Quick Start Kit (Beckman).

Validation of alternative splicing events

Total RNA was recovered from fungal cells grown in synthetic medium with or without exogenous thiamine (0 or 1 mg/L). RT-PCR was then carried out using SUPERSCRIPT IIITM reverse transcriptase and Phusion DNA polymerase. The cDNA fragment of CYP5076C1 was amplified by gene specific primers, primer-1 (5'-ATGG ATATCAAGGAAAAGCCGA-3'), primer-2 (5'-CTATA TAGCACGTTTTTGAAGTGTA-3'), and primer-3 (5'-AGGTCGGCAAGCTTGCG-3'). The reaction mixture (50 μ L) contained first-strand cDNA solution (1 μ L), dNTP (200 μ M), primers (2 μ M each), DMSO (2%), and Phusion DNA polymerase (0.02 U/ μ L) in Phusion HF buffer. The PCR products were separated by 1.5% agarose gel electro-

phoresis, stained with GelStar[®] Nucleic Acid Stain (TaKaRa), and visualized using Molecular Imager FX (Bio-Rad). cDNA fragments were purified, cloned into the pUC18 plasmid, and sequenced.

Sequence alignment and phylogenetic analysis

Multiple alignment of AoCYPs was carried out using the ClustalW program with a gap penalty of 10, a gap extension penalty of 0.2, and GONNET as protein matrix series (Thompson et al. 1994). The phylogenetic tree was constructed by the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) method with the Jones-Taylor-Thornton matrix using PHYLIP software (Felsenstein 1989), and visualized using the FigTree program.

Results and discussion

Genome-wide survey and molecular identification of AoCYPs

The filamentous fungus *A. oryzae* has eight chromosomes with an entire genome size of 37.6 Mb (Machida et al. 2005). The whole-genome sequence of *A. oryzae* strain RIB40 was released recently (<http://www.bio.nite.go.jp/dogan/Top>). According to the public database, there are several candidate genes assigned to P450s. However, some candidates have low sequence similarity to P450s but significantly higher similarity to other proteins, suggesting that annotational errors may be involved (data not shown). Several candidates have unexpected truncations of their N- and/or C-terminal sequence(s). Therefore, we further refined gene annotation accuracy based on the following sequence features: (1) conservation of F-x-x-G-x-x-x-C-x-G in the heme-binding domain, (2) conservation of E-x-x-R in the K-helix, (3) A/G-G-x-x-T at the center of the I-helix, and (4) a hydrophobic transmembrane domain (TMD) at the N-terminal region. After searching the database, 155 putative genes of P450 were identified from the whole-genome sequence (Fig. 1). However, sequence deletions and/or inframe stop codon(s) were found in 13 genes, suggesting that they are possibly pseudogenes that have originated from gene reorganizations and/or single mutations during fungal evolution (Fig. S1). Although some pseudogene-like AoCYPs were expressed, sequence deletions and/or inframe stop codon(s) were also verified from their transcripts by 3'- and 5'-RACE (Forhman 1993) and RT-PCR analyses (data not shown). Therefore, 142 AoCYPs from the *A. oryzae* genome were selected for further investigation.

According to the P450 nomenclature committee, families share greater than 40% identity, and subfamilies share

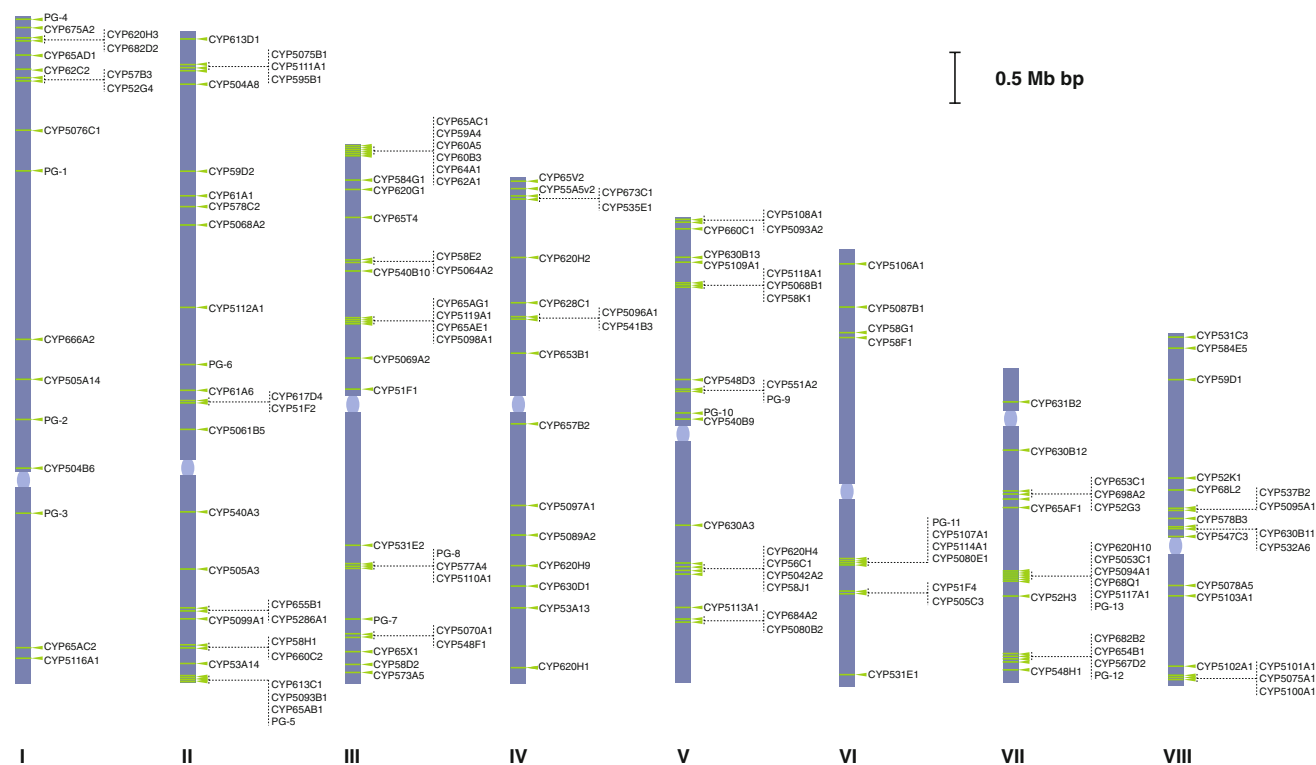


Fig. 1 Chromosomal localization of cytochrome P450 in *Aspergillus oryzae*. PG pseudogene

greater than 55% identity of amino acid sequences. The numbers following the root symbol CYP indicate the family and letters indicate the subfamily (Nelson et al. 1996). Based on sequence comparisons, 142 AoCYPs were assigned into 87 families. There were significantly more AoCYP genes than in other ascomycetous fungi, e.g., 122 in *Magnaporthe grisea*, 107 in *Fusarium graminearum*, 41 in *Neurospora crassa*, and 111 in *A. nidulans* (Deng et al. 2007; Kelly et al. 2008). A possible explanation for the marked increase in gene number may be that fungal P450s continuously diversified after separation of genera, and perhaps after speciation as well, indicating that sequences diversified within a short evolutionary period (Deng et al. 2007). During the evolutionary history of *Aspergillus* spp., *A. oryzae* would have vigorously expanded its genome size to gain extra genes by horizontal gene transfer and duplication. For instance, both gene number and genomic size are 25–30% larger in *A. oryzae* than in *A. nidulans* and *A. fumigatus*. The extra genes are likely to be involved in secondary metabolism (Machida et al. 2005). Therefore, the substantial increase in P450 gene number might indicate adaptation to the specific metabolic requirements of *A. oryzae* (<http://www.aspergillus.org.uk/index.html>). In fact, several AoCYPs were located near genes involved in biosynthesis of secondary metabolites, such as polyketide synthase and non-ribosomal peptide synthase. For example, CYP655B1 and CYP5286A1 are flanked by polyketide

synthase (protein ID; BAE56814.1) within 19 kbp distance on chromosome II, CYP5110A1 and CYP577A1 are flanked by polyketide synthase (protein ID; BAE58990.1) within 22 kbp distance on chromosome III, and CYP5099A1 is flanked by isopenicillin N synthase (protein ID; BAE56800.1) and non-ribosomal peptide synthase (protein ID; BAE56801.1) within 19 kbp distance on chromosome II. On the other hand, the genome of *A. oryzae* is strikingly similar to that of *A. flavus*; *A. flavus* has 12,197 genes in its 36.8 Mb genome (Payne et al. 2006). The presence of 159–167 *A. flavus* P450s has also been revealed by genomic annotation (Park et al. 2008; <http://drnelson.utm.edu/CytochromeP450.html>, <http://p450.riceblast.snu.ac.kr/index.php?a=view>). Sequence comparisons indicated that 138 of the P450s in *A. oryzae* and *A. flavus* show orthologous relationships, sharing amino acid sequence identities of more than 95%. This raises the question as to whether the P450s of both species exhibit similar transcriptional and functional profiles. In contrast, 16 species were found in *A. oryzae* but not in *A. flavus*. Some of these 16 AoCYPs were located within 20 kbp distance, suggesting that a cluster of secondary metabolism genes has developed specifically in *A. oryzae*. In fact, CYP5119A1, CYP65AG1, CYP65AE1, and CYP5098A1 are distributed within an 18-kbp region on chromosome III, and are flanked by several metabolic genes such as the non-ribosomal peptide synthase module (Protein ID; BAE60013.1), which is also

absent from *A. flavus*. Because CYP5119 is an orphan family that is only found in *A. oryzae*, the related gene cluster might play important roles in a secondary metabolic process that is unique to this fungus.

Insights into conserved sequences of AoCYPs

The protein architecture of P450s is generally well conserved, even though they show considerable sequence divergence (Graham and Peterson 1999). Structural conservation is presumably important for fundamental aspects of P450 activity, such as heme binding, acceptance of electrons, and activation of molecular oxygen. Therefore, the classical P450s contain conserved sequences that shape the core structure. The most characteristic P450 consensus sequence, F-x-x-G-x-x-x-C-x-G, is found in the heme-binding domain, where the conserved Cys serves as the fifth axial ligand to the heme. Consequently, we identified 142 AoCYPs-containing proximal Cys residues in the heme-binding domain. In addition, several possible functions of conserved amino acid residues have been proposed: (1) the first amino acid is Phe, whose phenyl group appears to protect the reactive cysteine ligand, (2) the fourth Gly appears to initiate the hairpin turn in the loop, and (3) the tenth Gly is small enough to come in close contact with the heme (Ortiz-de-Montellano 2005; Koymans et al. 1993). Amino acid substitutions in the heme-binding domain were identified from some AoCYPs, such as the first Phe to Trp/Tyr in 11 species and the tenth Gly to Ala in seven species; nevertheless, no significant detrimental effects on catalytic functions might be introduced due to their physicochemical characteristics similar to those in classical P450s. However, rare substitutions were found in several AoCYPs, such as the first Phe to Val (CYP5097A1) or Leu (CYP5116A1), and the fourth Gly to Ser (CYP660C1 and 660C2) or Trp (CYP5099A1), suggesting possible involvement of a unique and novel heme-binding feature in those AoCYPs. Moreover, a multiple alignment revealed that CYP5102A1 contained a substantially altered signature sequence, L-S-T-S-I-N-D-C-P-K. Although there is a paucity of literature on biochemical and functional characterization of fungal P450s, further research on AoCYPs could clarify the peculiar reaction mechanisms associated with unique sequences.

The core structure of the proximal side of P450s is additionally stabilized by a consensus sequence of ExxR in the K-helix and a coil known as the “meander” located between the K-helix and the heme-binding domain (Hasemann et al. 1995; Chen and Zhou 1992). The characteristic ExxR sequence is always conserved, and a possible “meander” sequence, normally PER, was found in each of the AoCYPs. In addition to the proximal side, many P450s have a conserved Thr residue in the distal I-helix that shapes the oxygen binding pocket, stabilizes the iron-oxo intermedi-

ate, and facilitates heterolytic cleavage of the O–O bond (Poulos et al. 1987; Koymans et al. 1993). A multiple alignment analysis revealed that 107 AoCYPs contain the Thr residue in the I-helix and nine AoCYPs substituted Thr to Ser at the appropriate position. It is possible that both OH-containing Thr and Ser exhibit the same function with regard to oxygen activation. However, neither Thr nor Ser was identified in the I-helix in 38 AoCYPs, suggesting possible involvements of unique reaction mechanisms. It has been known that some P450s such as plant allene oxide synthase (CYP74 family) and human prostaglandin I₂ synthase (CYP8 family), which lack conserved Thr, catalyze isomerization reactions of hydroperoxide compounds (Howe et al. 2000; Ullrich 2003). Thus, abnormal sequences in AoCYPs would be an attractive target to better understand the unique structural and mechanistic characteristics of P450s.

Membrane topology of AoCYPs

Most eukaryotic P450s are likely to have an N-terminal TMD sequence, which is responsible for subcellular localization to the endoplasmic reticulum (Nelson and Strobel 1988). A typical TMD contains 20–30 hydrophobic amino acid residues that shape a helical structure as a membrane anchor. The TMD-associated subcellular localization to membranes should also be important for protein–protein interactions with the membrane-anchored cytochrome P450 oxidoreductase (CPR), which is the common redox partner of eukaryotic P450s. A possible TMD was distinguished from 133 AoCYPs using SOSUI and TMHMM servers (Hirokawa et al. 1998; <http://bp.nuap.nagoya-u.ac.jp/sosui/>; <http://www.cbs.dtu.dk/services/TMHMM/>), suggesting that most AoCYPs localize to the endoplasmic reticulum. A well-known soluble nitric oxide reductase (P450_{nor}, CYP55A5) lacks the N-terminal TMD sequence (Nakahara et al. 1993). On the other hand, a distinctive N-terminal TMD was not found in several AoCYPs assigned to the P450 family CYP505 (CYP505A3, 505C3, and 505A14), CYP540 (CYP540A3, 540B9, and 540B10), CYP541 (CYP541B3), and CYP5053 (CYP5053C1). The CYP505 family contains self-sufficient P450s fused with a reductase domain, such as P450_{foxy}, which lacks a distinctive TMD but is capable of loosely binding to the membrane (Kitazume et al. 2000; Nakayama et al. 1996). This suggests that AoCYPs assigned to the CYP505 family might also be expressed in membrane fractions. The CYP541 and CYP540 families are phylogenetically close to the CYP505 family even though they lack the reductase domain. These families probably emerged *via* disconnection of an ancestral fusion P450. If they are not membrane-bound proteins, the membrane-anchored CPR should still weakly transfer electrons CYP540 and CYP541 families. In *A. oryzae*, we

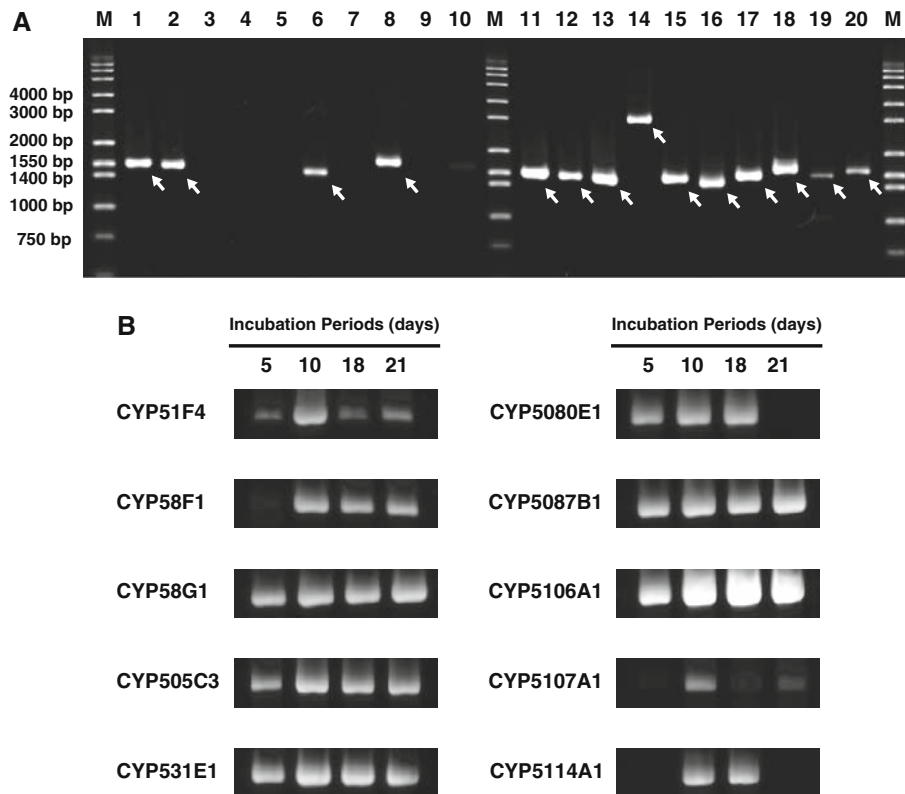


Fig. 2 Transcriptomic survey of AoCYPs encoded on chromosome VI. **a** Effects of culture conditions on gene expression of AoCYPs. RT-PCR was carried out using total RNA extracted from *A. oryzae* grown for 5 days in YPD liquid culture medium (Lanes 1–10) or 10 days in a nitrogen-limited synthetic liquid culture medium (Lanes 11–20). cDNA fragments of CYP51F4 (Lanes 1 and 11), CYP58F1 (Lanes 2 and 12), CYP58G1 Lane (Lanes 3 and 13), CYP505C3 (Lanes 4 and 14), CYP531E1 (Lanes 5 and 15), CYP5080E1 (Lanes 6 and 16), CYP5087B1 (Lanes 7 and 17), CYP5106A1 (Lanes 8 and 18), CYP5107A1 (Lanes 9 and 19), and CYP5114A1 (Lanes 10 and 20)

were separated by 1.5% agarose gel electrophoresis, stained with Gel-Star® Nucleic Acid Stain (TaKaRa), and visualized using Molecular Imager FX (BioRad). Lane M was loaded with a DNA marker. Arrows indicate amplified cDNA of AoCYPs. Primer sequences are listed in Supplemental Table. **b** Time course of gene expression in a nitrogen-limited synthetic liquid culture medium. RT-PCR was performed with total RNA individually extracted from *A. oryzae* grown for 5, 10, 18, or 21 days in a nitrogen-limited synthetic liquid culture medium. Analytical procedures were same as that for Fig. 2a

also annotated TMD-containing and TMD-lacking CPRs and evaluated their gene expression (Figs. S2, S3). The possession of both TMD-containing and TMD-lacking CPR would be advantageous to interact with various AoCYPs as a redox partner (Lah et al. 2008).

Transcriptomic survey of cytochromes P450 from *A. oryzae*

The current sequence database of P450s exponentially enlarged as a result of several genome projects (<http://drnelson.utmem.edu/CytochromeP450.html>; <http://p450.riceblast.snu.ac.kr/index.php?a=view>; Park et al. 2008). In addition to bioinformatic studies, experimental approaches are also necessary to evaluate practical applications in a post-genomic era. In this study, our principal aim was to isolate and characterize full-length cDNAs to clarify transcriptional capabilities of AoCYPs. Expression profiles of genes involved in secondary metabolic systems of *A. oryzae* are very likely to be affected by cultivation conditions

(Tamano et al. 2008; Machida et al. 2005). Therefore, we used YPD and synthetic liquid culture media for fungal growth (detailed in materials and methods). Figure 2 shows expression profiles of AoCYPs encoded on chromosome VI. When fungi were grown in YPD liquid culture medium, many genes were not amplified by RT-PCR, whereas some genes such as CYP51 and CYP58 families were confirmed to be expressed (Fig. 2a). No significant gene expression was also observed in a synthetic liquid culture medium under nitrogen-rich conditions (data not shown). P450s classified in CYP51 family are highly conserved across diverse organisms, and play important roles in steroid metabolism. Although CYP58 family in *Fusarium* spp. has been shown to be involved in biosynthesis of secondary metabolites, metabolic pathways associated with CYP58 family might also be important for fungal cells because a wide variety of fungal species possess homologous genes (Hohn et al. 1995; <http://drnelson.utmem.edu/CytochromeP450.html>; <http://p450.riceblast.snu.ac.kr/index.php?a=view>).

Table 1 The gene list of cytochrome P450 from *A. oryzae*

Chromosome Name	Accession number	Contig Location	NA (bp)	AA (residue)	Transcript	Chromosome Name	Accession number	Contig Location	NA (bp)	AA (residue)	Transcript		
I	CYP675A2	AB514764	SC005 7979–9791	1,527	508	FL	III	CYP573A5	AB514733	SC023 117944–116142	1,554	517	FL
	CYP620H3	AB514745	SC005 183457–181557	1,581	526	FL		CYP58D2	AB514691	SC023 193118–191351	1,548	515	FL
	CYP682D2	AB514766	SC005 187561–185597	1,548	515	FL		CYP65X1	AB514713	SC023 331674–333474	1,566	521	FL
	CYP65AD1	AB514708	SC005 302578–300796	1,509	502	FL		CYP548F1	AB514729	SC023 474339–472501	1,572	523	FL
	CYP62C2	AB514704	SC005 475900–477474	1,455	484	FL		CYP5070A1	AB514773	SC023 480458–482071	1,614	537	FL
	CYP57B3	AB514690	SC005 573238–575042	1,527	508	FL		CYP5110A1	AB514793	SC023 1141014–1139075	1,620	539	FL
	CYP52G4	AB514684	SC005 577203–575544	1,593	530	FL		CYP577A4	AB514734	SC023 1148972–1150552	1,530	509	FL
	CYP5076C1	AB514776	SC005 1087712–1085860	1,632	543	FL		CYP531E2	BAE59053.1	SC023 1342126–1344108	1,602	533	NE
	CYP666A2	BAE56143.1	SC005 3144893–3146865	1,791	596	FS		CYP65AC1	AB514706	SC026 11301–9652	1,530	509	FL
	CYP505A14	BAE56245.1	SC005 3516048–3519656	3,315	1,104	NE		CYP59A4	AB514698	SC026 59498–61206	1,593	530	FL
	CYP504B6	AB514717	SC005 4404626–4406412	1,578	525	FL		CYP60A5	BAE59521.1	SC026 63849–62258	1,488	495	NE
	CYP5116A1	AB514796	SC009 240943–242196	1,254	417	FL		CYP60B3	AB514700	SC026 65773–64217	1,503	500	FL
	CYP65AC2	AB514707	SC009 340207–341879	1,530	509	FL		CYP64A1	BAE59527.1	SC026 73168–75118	1,587	528	NE
	CYP65AB1	AB514705	SC001 63753–65572	1,473	490	FL		CYP62A1	AB514703	SC026 79811–78163	1,527	508	FL
	CYP5093B1	BAE56598.1	SC001 81343–78250	1,575	524	FS		CYP584G1	AB514738	SC026 270089–271778	1,521	506	FL
	CYP613C1	AB514740	SC001 90529–88933	1,500	499	FL		CYP620G1	AB514742	SC026 371119–373082	1,590	529	FL
	CYP53A14	BAE56637.1	SC001 182099–180290	1,602	533	NE		CYP65T4	AB514712	SC026 663809–662271	1,482	493	FL
	CYP660C2	AB514762	SC001 382760–381102	1,542	513	FL		CYP58E2	AB514692	SC026 1094103–1095957	1,536	511	FL
	CYP58H1	AB514695	SC001 383260–385028	1,533	510	FL		CYP5064A2	BAE59846.1	SC026 1096818–1098547	1,644	547	FS
	CYP5099A1	AB514786	SC001 636916–635420	1,497	498	FL		CYP540B10	AB514725	SC026 1149822–1148045	1,584	527	FL
	CYP5286A1	AB514800	SC001 708600–706690	1,527	508	FL		CYP65AG1	AB514711	SC026 1611203–1612908	1,524	507	FL
	CYP655B1	AB514759	SC001 718639–716987	1,524	507	FL		CYP5119A1	AB514799	SC026 1618599–1620164	1,566	521	FL
	CYP505A3	AB514718	SC001 1128777–1125241	3,312	1,103	FL		CYP65AE1	AB514709	SC026 1624148–1622434	1,539	512	FL
CYP540A3	BAE57130.1	SC001 1651339–1653041	1,653	550	NE		CYP5098A1	AB514785	SC026 1627559–1629097	1,539	512	FL	
CYP5061B5	AB514770	SC003 298711–300604	1,572	523	FL		CYP5069A2	BAE60165.1	SC026 2071209–2072992	1,581	526	FS	
CYP51F2	AB514681	SC003 549247–547657	1,524	507	FL		CYP51F1	AB514680	SC026 2309786–2311525	1,575	524	FL	
CYP617D4	AB514741	SC003 570250–568521	1,611	536	FL	IV	CYP620H1	AB514743	SC012 153166–151248	1,578	525	FL	
CYP61A6	AB514702	SC003 672417–674232	1,614	537	FL		CYP53A13	AB514687	SC012 765362–767112	1,554	517	FL	
CYP5112A1	BAE57727.1	SC003 1498000–1495794	1,578	525	FS		CYP630D1	AB514754	SC012 949158–950763	1,551	516	FL	
CYP5068A2	AB514771	SC003 2297146–2295650	1,497	498	FL		CYP620H9	AB514747	SC012 1160373–1162294	1,590	529	FL	
CYP578C2	AB514736	SC003 2467051–2465352	1,572	523	FL		CYP5089A2	AB514780	SC012 1434311–1432714	1,500	499	FL	
CYP61A1	AB514701	SC003 2585336–2583503	1,587	528	FL		CYP5097A1	AB514784	SC012 1742424–1743947	1,524	507	FL	
CYP59D2	AB514699	SC003 2825205–2823433	1,650	549	FL		CYP657B2	AB514760	SC012 2554858–2553078	1,725	574	FL	
CYP504A8	AB514716	SC003 3689601–3691331	1,551	516	FL		CYP620H2	AB514744	SC102 434114–436073	1,587	528	FL	

Table 1 continued

Chromosome Name	Accession number	Contig Location	NA (bp) AA (residue)	Transcript Chromosome Name	Accession number	Contig Location	NA (bp) AA (residue)	Transcript					
IV	CYP595B1	AB514739	SC003 3811869–3813420	1,500	499	FL	CYP628C1	AB514749	SC102 892232–890247	1,608	535	FL	
	CYP5111A1	AB514794	SC003 3848877–3850649	1,542	513	FL	CYP5096A1	BAE61439.1	SC102 1050720–1052486	1,521	506	FS	
	CYP5075B1	AB514775	SC003 3851245–3853084	1,671	556	FL	CYP541B3	AB514726	SC102 1054010–1055472	1,404	467	FL	
	CYP613D1	BAE58578.1	SC003 4127673–4125910	1,533	510	NE	CYP653B1	AB514756	SC102 1409731–1411563	1,533	510	FL	
	CYP65V2	BAE61720.1	SC166 198122–199828	1,524	507	FS	CYP548H1	AB514730	SC011 151615–149759	1,560	519	FL	
	CYP55A5v2	AB514688	SC166 201149–202795	1,227	408	FL	CYP57D2	AB514732	SC011 286384–284536	1,440	479	FL	
	CYP673C1	AB514763	SC166 282991–284716	1,569	522	FL	CYP654B1	AB514758	SC011 286826–288616	1,560	519	FL	
	CYP535E1	AB514722	SC166 312204–310498	1,539	512	FL	CYP682B2	AB514765	SC011 291007–288916	1,566	521	FL	
	V	CYP551A2	AB514731	SC113 145541–147583	1,524	507	FL	CYP52H3	AB514685	SC011 887837–886117	1,557	518	FL
		CYP548D3	AB514728	SC113 252847–254610	1,593	530	FL	CYP5117A1	AB514797	SC011 1051191–1049375	1,521	506	FL
		CYP58K1	AB514697	SC113 1184864–1183051	1,536	511	FL	CYP68Q1	AB514715	SC011 1051900–1053759	1,647	548	FL
		CYP5068B1	AB514772	SC113 1193673–1192063	1,611	536	FL	CYP5053C1	AB514769	SC011 1099331–1097727	1,410	469	FL
CYP5118A1		AB514798	SC113 1197216–1195543	1,497	498	FL	CYP620H10	AB514748	SC011 1103282–1105268	1,587	528	FL	
CYP5109A1		AB514792	SC113 1411903–1413731	1,503	500	FL	CYP65AF1	AB514710	SC011 1761174–1762901	1,539	512	FL	
CYP630B13		AB514753	SC113 1451229–1449585	1,578	525	FL	CYP52G3	AB514683	SC011 1822625–1824269	1,557	518	FL	
CYP660C1		AB514761	SC113 1760438–1762088	1,545	514	FL	CYP698A2	BAE65190.1	SC011 1893711–1891669	1,548	515	FS	
CYP5093A2		AB514781	SC113 1807380–1809053	1,674	557	FL	CYP653C1	AB514757	SC011 1897240–1899079	1,560	519	FL	
CYP5108A1		BAE63248.1	SC113 1834519–1832963	1,449	482	NE	CYP630B12	AB514752	SC011 2373084–2374733	1,587	528	FL	
CYP540B9		AB514724	SC124 135791–137597	1,680	559	FL	CYP631B2	AB514755	SC206 200368–202191	1,533	510	FL	
CYP630A3		AB514750	SC111 740134–738572	1,563	520	FL	VIII	CYP531C3	AB514719	SC010 37762–36045	1,575	524	FL
CYP620H4	AB514746	SC111 1113425–1111497	1,581	526	FL	CYP584E5		AB514737	SC010 183805–182249	1,506	501	FL	
CYP56C1	AB514689	SC111 1160825–1162558	1,488	495	FL	CYP59D1		BAE66084.1	SC010 470087–471859	1,647	548	NE	
CYP5042A2	AB514768	SC111 1201326–1203015	1,593	530	FL	CYP52K1		AB514686	SC010 1457139–1459052	1,584	527	FL	
CYP58J1	AB514696	SC111 1237470–1239250	1,530	509	FL	CYP68L2		AB514714	SC010 1546664–1544898	1,524	507	FL	
CYP684A2	AB514767	SC111 1694844–1693125	1,467	488	FL	CYP537B2		AB514723	SC010 1749258–1747595	1,476	491	FL	
CYP5080B2	AB514778	SC111 1717617–1719320	1,518	505	FL	CYP5095A1		AB514783	SC010 1762628–1760843	1,560	519	FL	
CYP5113A1	AB514795	SC111 1,581541–1583308	1,554	517	FL	CYP578B3		AB514735	SC010 1822486–1824059	1,497	498	FL	
VI	CYP5107A1	AB514791	SC020 601589–603342	1,575	524	FL		CYP630B11	AB514751	SC010 1890353–1888785	1,569	522	FL
	CYP5114A1	BAE63475.1	SC020 657999–656189	1,545	514	FS		CYP532A6	AB514721	SC010 1903891–1902219	1,560	519	FL
	CYP5080E1	BAE63486.1	SC020 691554–689805	1,506	501	FS		CYP547C3	AB514727	SC010 1953242–1955042	1,641	546	FL
	CYP51F4	AB514682	SC020 885080–883469	1,542	513	FL		CYP5078A5	AB514777	SC103 365151–366718	1,521	506	FL
	CYP505C3	BAE63564.1	SC020 916089–912562	3,156	1051	FS	CYP5103A1	BAE65635.1	SC103 463173–464998	1,503	500	FS	

Table 1 continued

Chromosome Name	Accession number	Contig	Location	NA (bp)	AA (residue)	Transcript	Chromosome Name	Accession number	Contig	Location	NA (bp)	AA (residue)	Transcript
CYP531E1	AB514720	SC020	1735772–1734027	1,530	509	FL	CYP5102A1	AB514789	SC103	1128173–1129610	1,338	445	FL
CYP5087B1	AB514779	SC038	241250–239612	1,533	510	FL	CYP5101A1	AB514788	SC103	1243582–1241748	1,641	546	FL
CYP58G1	AB514694	SC038	500696–498976	1,497	498	FL	CYP5075A1	AB514774	SC103	1244215–1245996	1,659	552	FL
CYP58F1	AB514693	SC038	531961–530249	1,503	500	FL	CYP5100A1	AB514787	SC103	1252988–1254893	1,722	573	FL
CYP5106A1	AB514790	SC138	137250–139040	1,605	534	FL							

The name of CYP was assigned by P450 nomenclature committee. AoCYPs isolated as full-length cDNA were listed with accession number. Non-expressed and frame-shifted AoCYPs were listed with protein ID assigned for predicted sequence (<http://www.bio.nite.go.jp/dogan/Top>). Contigs are listed at DOGAN (<http://www.bio.nite.go.jp/dogan/Top>). NA and AA indicate length of nucleic acid of cDNA and deduced amino acid, respectively. The type of transcript was indicated as FL for full-length cDNA, FS for frame-shifted cDNA, and NE for not-expressed gene

Thus, we expected that AoCYPs involved in housekeeping pathways would be expressed in YPD and nitrogen-rich synthetic culture media. In contrast, a series of AoCYP genes were strongly expressed when *A. oryzae* was grown in a synthetic liquid culture medium under nitrogen-limited conditions (Fig. 2a), suggesting that transcriptional regulation of AoCYP responds to nitrogen limitation or starvation. Previously, P450-dependent metabolic pathways in white-rot basidiomycetes such as *Phanerochaete chrysosporium* and *Coriolus versicolor* have been shown to be activated under nitrogen-limited conditions (Ichinose et al. 1999; Matsuzaki and Wariishi 2004). Thus, there may be a unique mechanism that activates the fungal secondary metabolic system during nitrogen limitations. Although AoCYPs encoded on chromosome VI showed different time course of gene expression, a significant expression level of AoCYPs appeared after 5 days incubation and continued until 21 days incubation (Fig. 2b). Therefore, amplification and isolation of cDNAs by RT-PCR were carried out using a RNA cocktail which was prepared by mixing total RNA obtained from 5-, 10-, 18-, and 21-day-old mycelia grown in a synthetic liquid culture medium under nitrogen-limited conditions.

Using an RT-PCR technique, we determined transcriptional capabilities of 133 AoCYPs experimentally (Table 1). To our knowledge, this is the first report of experimental validation of AoCYPs expression, and our results provide evidence that a series of P450s can be expressed in ascomycetous fungi. So far, we isolated 121 full-length cDNAs encoding a mature open reading frame. Identification of these clones will be an advantage for generating recombinant systems, which can contribute to characterization and practical applications of AoCYPs. In addition, the experimentally deduced sequences will improve bioinformatic algorithms; in fact, we identified several mature AoCYPs with novel intron/exon boundaries, which were unexpected and miss-annotated in the database (<http://www.bio.nite.go.jp/dogan/Top>). The isolated and predicted cDNA sequences of 142 AoCYPs are listed in Figures S2 and S3. Interestingly, our results showed that several AoCYPs were alternatively spliced in response to different culture conditions. For example, CYP5076C1 was spliced to produce frame-shifted variants when *A. oryzae* was grown with exogenously added thiamine, while it was differently spliced to produce a mature variant when grown without thiamine (Fig. 3). Although further investigations are required to confirm whether exogenous thiamine directly affects splicing events, our results strongly suggest that unique splicing mechanisms, such as riboswitching, might be involved. In our experimental conditions, we isolated 12 immature AoCYPs whose open reading frames were shifted by illegal splicing events. Their functional expression might also be regulated by sophisticated maturation

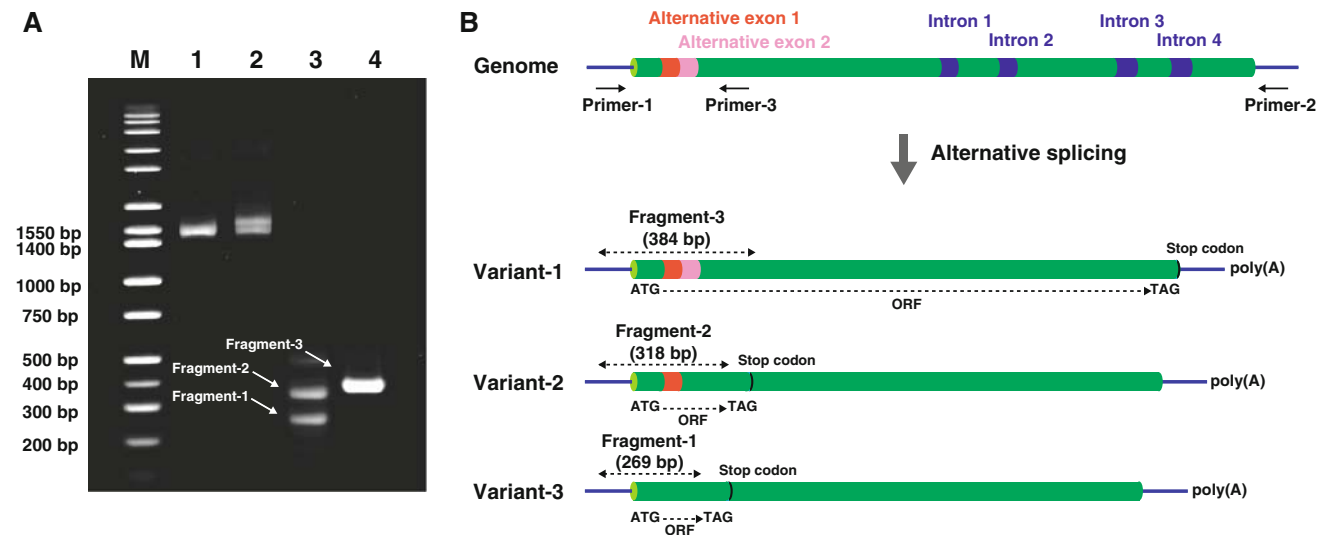


Fig. 3 Thiamine-dependent alternative splicing of AoCYP. **a** Alternative splicing of CYP5076C1 analyzed by RT-PCR. RT-PCR was carried out using total RNA extracted from *A. oryzae* grown for 18 days in a thiamin-containing synthetic liquid culture medium (Lanes 1 and 3) or thiamin-free synthetic liquid culture medium (Lanes 2 and 4). Lanes 1 and 2 were loaded with PCR products amplified by primer-1

and primer-2. Lanes 3 and 4 were loaded with PCR products amplified by primer-1 and primer-3. Fragments 1, 2, and 3 are illustrated in Fig. 3b. **b** Variant-1 was obtained from thiamin-free media. Variant-2 and variant-3 were obtained from thiamin-containing media. Columns indicate exons and introns. Lines indicate untranslated region. ORF indicates open reading frame

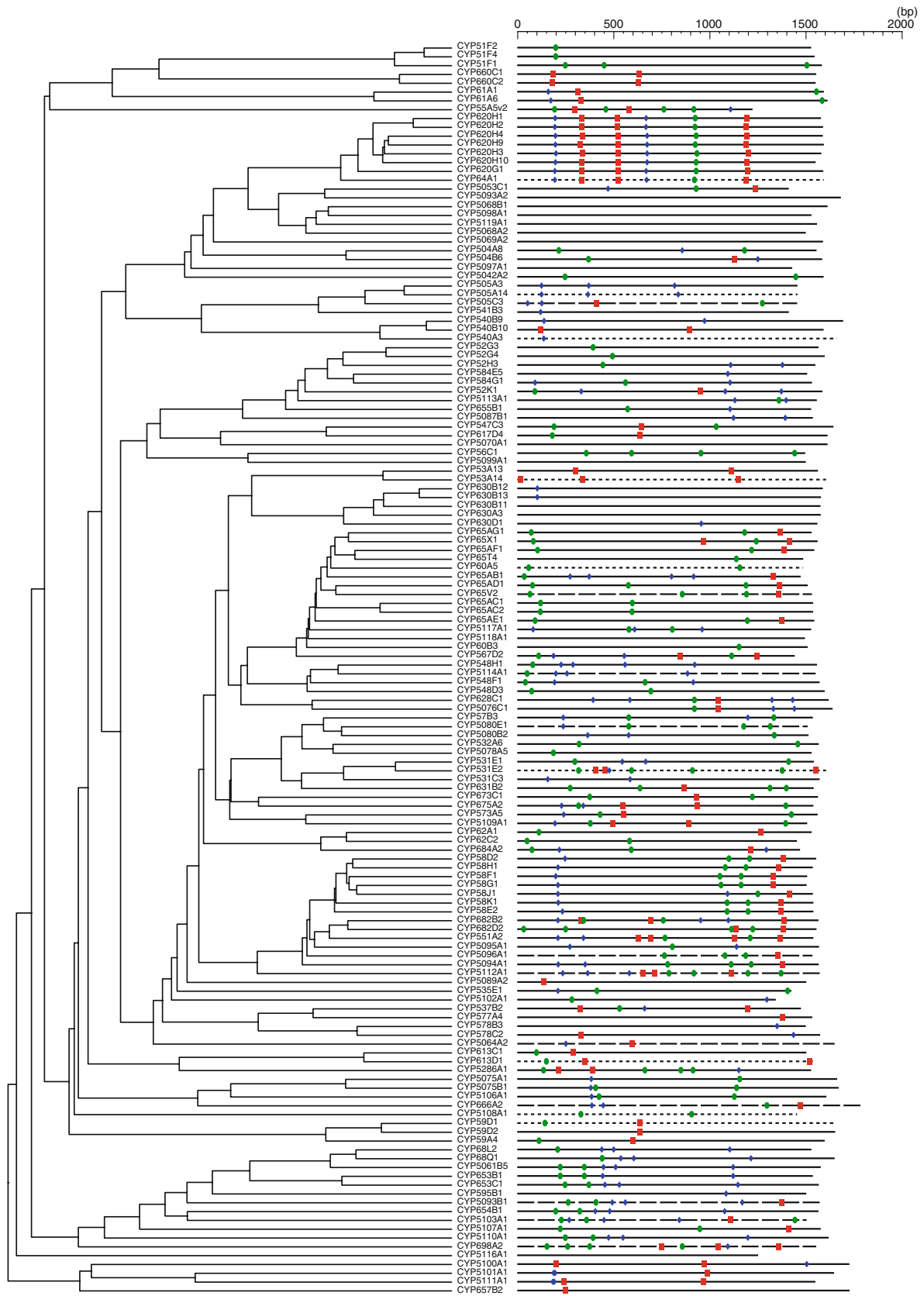
mechanisms at a post-transcriptional stage (Winkler et al. 2002; Cheah et al. 2002; Kubodera et al. 2003; Thore et al. 2006). Thus, a combination of quantitative and qualitative transcriptional profiling of AoCYPs is very important to understand physiological impacts on fungal metabolic activities.

Several *Aspergillus* spp. are known to produce polyketide derivatives, i.e., aflatoxins, which have serious toxic, mutagenic, and carcinogenic activities. However, the non-aflatoxigenic status of *A. oryzae* has been firmly established, because the aflatoxin biosynthesis pathway is inactivated. CYP64A1, a gene homolog involved in the aflatoxin biosynthesis pathway in *A. flavus* (Prieto and Woloshuk 1997), was transcriptionally silent in the synthetic culture medium in which a number of AoCYPs were expressed. This result is consistent with the inability of the fungus to produce aflatoxin. However, *A. oryzae* expressed seven AoCYPs assigned to the CYP620 family, which is phylogenetically close to the CYP64 family, suggesting a possible involvement of another polyketide biosynthesis pathway. In terms of human health and economically important products, polyketides have both positive (e.g., antibiotics and cancer therapeutic drug) and negative (e.g. mycotoxin) effects. Therefore, further investigations on their catalytic functions would be of great interest to better understand the safety and capabilities of this fungus (Barbesgaard et al. 1992; Machida et al. 2008).

Phylogeny and gene structure of cytochromes P450 in *A. oryzae*

Evolutionary histories of eukaryotic genes involve various trajectories such as gains and losses of introns. Although mechanisms and contributions of intron gain/loss events remain elusive, fossil aspects of introns can be helpful to unravel the dynamics of gene evolution. Figure 4 shows the phylogenetic tree and intron–exon organization of AoCYPs accurately constructed with experimentally deduced sequences. Multiple alignments of the deduced sequences and intron–exon structures revealed phylogenetic diversity of AoCYPs. Intron–exon organizations of P450 genes are generally conserved in plants, animals, and basidiomycetous fungi (Doddapaneni et al. 2005; Paquette et al. 2000; Tijet et al. 2001). The extremely diverse gene structure of AoCYPs might indicate that AoCYPs have emerged from a number of parent genes in the fungal ancestor (Deng et al.

Fig. 4 Phylogenetic tree and gene structures of AoCYPs. Phylogenetic tree was constructed by UPGMA methods using experimentally deduced sequences of isolated AoCYPs and bioinformatically predicted sequences for non-isolated cDNAs. Gene structures of AoCYPs were illustrated by solid line for full-length gene, dashed line for frame-shifted gene, and dotted line indicates non-expressed gene. Green circle, blue diamond, and red square indicate phase-0 intron, phase-1 intron, and phase-2 intron, respectively. Self-sufficient CYP505A3, 505A14, and 505C3 show catalytic P450 domain but not reductase domain



0.3

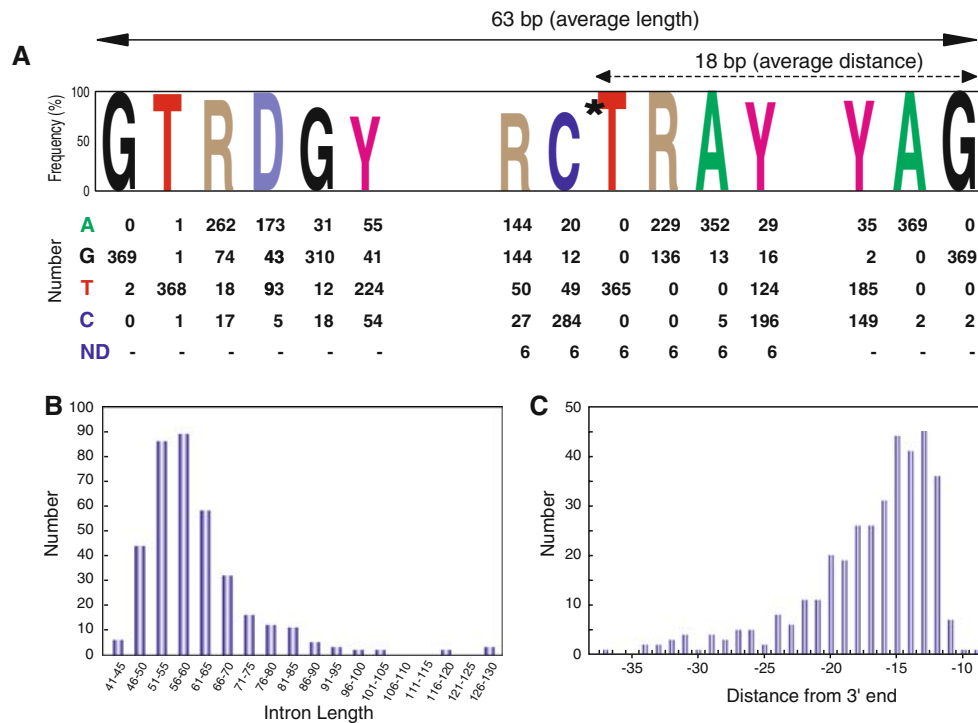


Fig. 5 Sequence conservation and molecular aspects of AoCYP introns. **a** Consensus sequences around 5'- and 3'-ends of intron. A, G, T, and C indicate adenosine, guanosine, thymidine, and cytidine, respectively. R and Y indicate purine nucleosides (A or G) or pyrimidine nucleosides (G or T), respectively. D indicates A, G, or T. Asterisk indicates conserved T in lariat sequence. Undiscriminated nucleotides for lariat sequences are listed as ND. **b** Distribution of introns length. **c** Distribution of distance for conserved T in lariat sequence

2007). The phylogenetic analysis also suggested an evolutionary trajectory in which gene duplication events were restricted in *A. oryzae*. This suggests that molecular mechanisms such as repeat-induced point mutations (RIP) may be involved (Galagan et al. 2003; Ikeda et al. 2002; Montiel et al. 2006). However, it appeared that several CYP families were generated or enlarged by evolutionary duplication, and some AoCYPs in such families were not expressed. Although RIP-like phenomena in *A. oryzae* are poorly understood, we can suggest a hypothetical scenario in which some AoCYPs are transcriptionally silenced by RIP-like mechanisms. For instance, the considerable similarities of sequences and gene structures between CYP64 and CYP620 families indicate evolutionary duplication events; however, CYP64A1 was not expressed under our experimental conditions, whereas CYP620 was abundantly expressed (Fig. 4). Recently, clan-level classification has been proposed for a higher order grouping of P450 families. The main concept of this analysis is that genes within a clan share a common ancestor gene and catalytic functions (Nelson 1998, 1999). Although common parameters for clan membership have not been clearly defined, CYP64 and CYP620 families would be classified into the same clan because these families were branched on a neighbor-joining tree with bootstrap values >70% (Deng et al. 2007). Since the isolated cDNAs of AoCYPs would be a powerful

advantage to facilitate downstream applications such as functional characterization using recombinant enzymes, further research on AoCYPs should aim to clarify the relationships between phylogeny and functions.

To better understand the molecular aspects of AoCYPs, we analyzed 371 intronic sequences in 121 experimentally validated AoCYPs. The average number of introns was 3.3 in each AoCYP gene, which is higher than the overall average 1.9 among all *A. oryzae* genes (Wang et al. 2008). The average intron length was 63 bp, twofold shorter than the overall average length among all *A. oryzae* genes (Wang et al. 2008). These data suggest that AoCYPs genes were organized more vigorously than other genes in *A. oryzae*.

The vast majority of introns conserved the dinucleotide GT at their 5'-end and AG at their 3'-end (Fig. 5a). In addition, the penultimate position from the 3'-end of introns was a pyrimidine nucleotide, and there was significant nucleotide consensus at the 5'-end. Furthermore, 365 introns encoded a characteristic lariat sequence in which the conserved T was usually located at the 18 position. Although the consensus sequences of introns are important for RNA maturation, some peculiar introns such as TA-CC and TG-CC in CYP540B10 and GC-AG in CYP620H9 were identified by comparing sequences of isolated cDNAs with those in genomic database. To the best of our knowledge, this is the first report describing TA-CC and TG-CC in introns from

any organisms, whereas GC-AG introns are rarely found in several other organisms (Rep et al. 2006). As shown in Fig. 5b and c, intron length and lengths of lariat sequences showed clear distributions. This characteristic is potentially useful for identification of novel P450s in ascomycetous fungi. Besides intronic sequences, microexons consisting of 6 or 10 nucleotides were found in CYP5111A1, CYP682B2, and CYP5101A1. The lengths of introns flanking microexons (52–75) were close to average size, suggesting that these unique gene structures were probably generated by stepwise insertion of two introns.

In conclusion, this study describes the molecular diversity of AoCYPs, investigated using experimental and bioinformatic approaches. The experimentally validated sequences and gene structures will enable molecular localization and characterization of novel P450s from ascomycetous fungi. Although further investigations are required to better understand transcriptional and post-transcriptional regulation of AoCYPs, it is clear that sophisticated molecular mechanisms enable superior metabolic performance of *A. oryzae*. To date, only a few fungal P450s have been functionally characterized. The isolated cDNAs will be useful in advanced studies on functional surveys of AoCYPs using recombinant systems. Potential practical applications of AoCYP will be explored in the near future.

Acknowledgments We are grateful to Hiroko Tsutsumi and Yoji Hata (Res. Inst., Gekkeikan Sake Co.) for providing *A. oryzae* strain RIB40 and for valuable technical advices. This research was supported in part by an NIRS Research Grant (Noda Institute for Scientific Research; to H. I.).

References

- Abe K, Gomi K, Hasegawa F, Machida M (2006) Impact of *Aspergillus oryzae* genomics on industrial production of metabolites. *Mycopathologia* 162:143–153
- Aoyama Y, Noshiro M, Gotoh O, Imaoka S, Funae Y, Kurosawa N, Horiuchi T, Yoshida Y (1996) Sterol 14-demethylase P450 (P45014DM) is one of the most ancient and conserved P450 species. *J Biochem* 119:926–933
- Barbesgaard P, Heldt-Hansen HP, Diderichsen B (1992) On the safety of *Aspergillus oryzae*: a review. *Appl Microbiol Biotechnol* 36:569–572
- Cheah MT, Wachter A, Sudarsan N, Breaker R (2002) Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. *Nature* 447:497–500
- Chen S, Zhou D (1992) Functional domains of aromatase cytochrome P450 inferred from comparative analyses of amino acid sequences and substantiated by site-directed mutagenesis experiments. *J Biol Chem* 267:22587–22594
- Demain AL, Fang A (2000) The natural functions of secondary metabolites. *Adv Biochem Eng Biotechnol* 69:1–39
- Deng J, Carbone I, Dean RA (2007) The evolutionary history of cytochrome P450 genes in four filamentous Ascomycetes. *BMC Evol Biol* 7:30
- Doddapaneni H, Chakraborty R, Yadav JS (2005) Genome-wide structural and evolutionary analysis of the P450 monooxygenase genes (P450ome) in the white rot fungus *Phanerochaete chrysosporium*: evidence for gene duplications and extensive gene clustering. *BMC Genomics* 6:92
- Felsenstein J (1989) PHYLIP—phylogeny inference package (version 3.2). *Cladistics* 5:164–166
- Forhman MA (1993) Rapid amplification of complementary DNA ends for generation of full-length complementary DNAs: thermal RACE. *Methods Enzymol* 218:340–356
- Galagan JE, Calvo SE, Borkovich KA et al (2003) The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* 422:859–868
- Galagan JE, Calvo SE, Cuomo C et al (2005) Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* 438:1105–1115
- Gotoh O (1993) Evolution and differentiation of P450 genes. In: Omura T, Ishimura Y, Fujii-Kuriyama Y (eds) *Cytochrome P-450*, 2nd edn. Kodansha, Tokyo, pp 255–272
- Graham SE, Peterson JA (1999) How similar are P450s and what can their differences teach us. *Arch Biophys Biochem* 369:24–29
- Hasemann CA, Ravichandran KG, Boddupalli SS, Peterson JA, Deisenhofer J (1995) Structure and function of cytochrome P450: a comparative analysis of three crystal structures. *Structure* 3:41–62
- Hirokawa T, Boon-Chiang S, Mitaku S (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* 14:378–379
- Hohn TM, Desjardins AE, McCormick SP (1995) The Tri4 gene of *Fusarium sporotrichioides* encodes a cytochrome P450 monooxygenase involved in trichothecene biosynthesis. *Mol Gen Genet* 248:95–102
- Howe GA, Lee GI, Itoh A, Li L, DeRocher AE (2000) Cytochrome P450-dependent metabolism of oxylipins in tomato. Cloning and expression of allene oxide synthase and fatty acid hydroperoxide lyase. *Plant Physiol* 123:711–724
- Ichinose H, Wariishi H, Tanaka H (1999) Biotransformation of recalcitrant 4-methylthiophene to water-extractable products using lignin-degrading basidiomycete *Coriolus versicolor*. *Biotechnol Prog* 15:706–714
- Ikeda K, Nakayashiki H, Kataoka T, Tamba T, Hashimoto Y, Tosa Y, Mayama S (2002) Repeat-induced point mutation (RIP) in *Magnaporthe grisea*: implications for its sexual cycle in the natural field context. *Mol Microbiol* 45:1355–1364
- Intikhab A, Hubbard SJ, Oliver SG, Rattray M (2007) A kingdom-specific protein domain HMM library for improved annotation of fungal genomes. *BMC Genomics* 8:97
- Kelly DE, Kraevac N, Mullins J, Nelson DR (2008) The CYPome (Cytochrome P450 complement) of *Aspergillus nidulans*. *Fung Genet Biol* 46:S53–S61
- Kirk TK, Schultz E, Connors WJ, Lorenz LF, Zeikus JG (1978) Influence of culture parameters on lignin metabolism by *Phanerochaete chrysosporium*. *Arch Microbiol* 117:277–285
- Kitazume T, Takaya N, Nakayama N, Shoun H (2000) *Fusarium oxysporum* fatty-acid subterminal hydroxylase (CYP505) is a membrane-bound eukaryotic counterpart of *Bacillus megaterium* cytochrome P450BM3. *J Biol Chem* 275:39734–39740
- Koymans L, Donne-Op Den Kelder GM, Koppele Te JM, Vermeulen NPE (1993) Cytochromes P450: their active site structure and mechanism of oxidation. *Drug Metab Rev* 25:325–387
- Kubodera T, Watanabe M, Yoshiuchi K, Yamashita N, Nishimura A, Nakai S, Gomi K, Hanamoto H (2003) Thiamine-regulated gene expression of *Aspergillus oryzae* thiA requires splicing of the intron containing a riboswitch-like domain in the 5'-UTR. *FEBS Lett* 555:516–520
- Lah L, Krasevec N, Trontelj P, Komel R (2008) High diversity and complex evolution of fungal cytochrome P450 reductase: cytochrome P450 systems. *Fung Genet Biol* 45:446–458

- Lewis DFL, Watson E, Lake BG (1998) Evolution of the cytochrome P450 superfamily: sequence alignments and pharmacogenetics. *Mutat Res* 410:245–270
- Machida M, Asai K, Sano M et al (2005) Genome sequencing and analysis of *Aspergillus oryzae*. *Nature* 438:1157–1161
- Machida M, Yamada O, Gomi K (2008) Genomics of *Aspergillus oryzae*: learning from the history of koji mold and exploration of its future. *DNA Res* 15:173–183
- Matsuzaki F, Wariishi H (2004) Functional diversity of cytochrome P450s of the white-rot fungus *Phanerochaete chrysosporium*. *Biochem Biophys Res Commun* 324:387–393
- Montiel MD, Lee HA, Archer DB (2006) Evidence of RIP (repeat-induced point mutation) in transposase sequence of *Aspergillus oryzae*. *Fung Genet Biol* 43:439–445
- Nakahara K, Hatano K, Usuda K, Shoun H (1993) Cytochrome P-450 55A1 (P-450dNIR) acts as nitric oxide reductase employing NADH as the direct electron donor. *J Biol Chem* 268:8350–8355
- Nakayama N, Takemae A, Shoun H (1996) Cytochrome P450foxy, a catalytically self-sufficient fatty acid hydroxylase of the *Fusarium oxysporum*. *J Biochem* 119:435–440
- Nelson DR (1998) Metazoan cytochrome P450 evolution. *Comp Biochem Physiol C Pharmacol Toxicol Endocrinol* 121:15–22
- Nelson DR (1999) Cytochrome P450 and the individuality of species. *Arch Biochem Biophys* 369:1–10
- Nelson DR, Strobel HW (1988) On the membrane topology of vertebrate cytochrome P450 proteins. *J Biol Chem* 263:6038–6050
- Nelson DR, Kamataki T, Waxman DJ et al (1993) The P450 superfamily: update on new sequences, gene mapping, accession numbers, early trivial names of enzymes and nomenclature. *DNA Cell Biol* 12:1–51
- Nelson DR, Koymans L, Kamataki T et al (1996) P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenet* 6:1–42
- Ortiz-de-Montellano PR (2005) Cytochrome P-450: structure, mechanism, and biochemistry, 3rd edn. Kluwer Academic/Plenum Publishers, New York
- Paquette SM, Bak S, Feyereisen R (2000) Intron-exon organization and phylogeny in a large superfamily, the paralogous cytochrome P450 genes of *Arabidopsis thaliana*. *DNA Cell Biol* 19:307–317
- Park J, Lee S, Choi J, Ahn K, Park B, Park J, Kang S, Lee YH (2008) Fungal cytochrome P450 database. *BMC Genomics* 9:402
- Payne GA, Nierman WC, Wortman JR, Pritchard BL, Brown D, Dean RA, Bhatnagar D, Cleveland TE, Machida M (2006) Whole genome comparison of *Aspergillus flavus* and *A. oryzae*. *Med Mycol* 44:9–11
- Poulos TL, Finzel BC, Haward AJ (1987) High-resolution crystal structure of P450cam. *J Mol Biol* 195:687–700
- Prieto R, Woloshuk CP (1997) *ord1*, an oxidoreductase gene responsible for conversion of O-methylsterigmatocytin to aflatoxin in *Aspergillus flavus*. *Appl Environ Microbiol* 63:1661–1666
- Rep M, Duyvesteijn RGE, Gale L, Usgaard T, Cornelissen BJC, Ma L-J, Ward TJ (2006) The presence of GC-AG introns in *Neurospora crassa* and other euascomycetes determined from analysis of complete genomes: implications for automated gene prediction. *Genomics* 87:338–347
- Tailor MJ, Richardson T (1979) Applications of microbial enzymes in food systems and in biotechnology. *Adv Appl Microbiol* 25:7–35
- Tamano K, Sano M, Yamane N, Terabayashi Y, Toda T, Sungawa M, Koike H, Hatamoto O, Umitsuki G, Takahashi T, Koyama Y, Asai R, Abe K, Machida M (2008) Transcriptional regulation of genes on the non-syntenic blocks of *Aspergillus oryzae* and its functional relationship to solid-state cultivation. *Fung Genet Biol* 45:139–151
- Thompson JD, Higgins DG, Gibson TJ (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
- Thore S, Leibundgut M, Ban N (2006) Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand. *Science* 312:1208–1211
- Tijet N, Helvig C, Feyereisen R (2001) The cytochrome P450 gene superfamily in *Drosophila melanogaster*: annotation, intron-exon organization and phylogeny. *Gene* 262:189–198
- Ullrich V (2003) Thoughts on thiolate tethering. Tribute and thanks to a teacher. *Arch Biochem Biophys* 409:45–51
- Wang K, Ussery DW, Brunak S (2008) Analysis and prediction of gene splice sites in four *Aspergillus* genomes. *Fung Genet Biol* 46:S14–S18
- Winkler W, Nahvi A, Breaker RR (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* 419:952–956
- Yoshida Y (1993) Sterol biosynthesis. In: Omura T, Ishimura Y, Fujii-Kuriyama Y (eds) *Cytochrome P-450*, 2nd edn. Kodansha, Tokyo, pp 93–101