Isolation and functional characterization of two rapeseed genes orthologous to Arabidopsis thaliana Phenylalanine Ammonia-lyase 2

Yu NI1,†, Yan-Zhen DONG2,†, Guan-Rong LI1, Bo LEI1, You-Rong CHAI1,*, Jia-Na LI1,*

1Chongqing Rapeseed Engineering Research Center; Chongqing Key Laboratory of Crop Quality Improvement; Key Laboratory of Biotechnology and Crop Quality Improvement of Ministry of Agriculture; College of Agronomy and Biotechnology, Southwest University, Tiansheng Road 216#, Beibei, Chongqing 400716, P. R. - CHINA
2Animal Science Department, Xichang College (North section), Xichang, Sichuan 615013, P. R. - CHINA

Received: 10.10.2008

Abstract: Two rapeseed (Brassica napus L.) phenylalanine ammonia-lyase 2 genes (BnPAL2-1 and BnPAL2-2) were isolated. The full-length cDNAs of BnPAL2-1 and BnPAL2-2 are 2422 bp and 2421 bp, with corresponding genomic sequences of 2735 bp and 2742 bp, respectively. Clues from NCBI BLAST, multi-alignment, the phylogenetic tree, conserved domains, preferential amino acids, and tertiary structures all indicated that they are orthologs of Arabidopsis thaliana PAL2 (AtPAL2). The purified His-tagged BnPAL2-1 and BnPAL2-2 proteins, expressed in Escherichia coli, were both active for PAL catabolism, but BnPAL2-2 was much more active than BnPAL2-1. Southern blot analysis showed that the BnPAL2 subfamily might contain 2-6 members. Transcripts of both BnPAL2-1 and BnPAL2-2 were detectable in all tested organs, but BnPAL2-2 was more organ-specific than BnPAL2-1. Transcription of BnPAL2-1 was most abundant in buds, with sequentially decreased abundance in stems, roots, flowers, seeds of all stages, silique pericarps and leaves, and was weak in hypocotyls and cotyledons. Expression of BnPAL2-2 was most intensive in roots, declining sequentially in buds, stems, flowers, and 30D seeds, and was weak in all other organs. Transcription of both genes in yellow-seeded rapeseed line L2’s most reproductive organs, including developing seeds, was lower than that of black-seeded line L1. These results, combined with our previous investigations on BnPAL1 genes, demonstrate that most Brassicaceae species contain orthologs of all 4 AtPAL genes. Both PAL1 and PAL2 were most probably triplicated in Brassicaceae ancestor and simultaneous down-regulation of all/most PAL paralogs contributes to the formation of the yellow seed trait in L2.

Key words: Arabidopsis thaliana, Brassica napus L., ortholog, phenylalanine ammonia-lyase (PAL), yellow seed

Abbreviations: D: Days after flowering; PAL: phenylalanine ammonia-lyase; PAGE: polyacrylamide gel electrophoresis; RACE: rapid amplification of cDNA ends; RT: reverse transcription; SDS: sodium dodecyl sulfate; UTR: untranslated region.

Introduction

Phenylpropanoids are essential for growth, development, and survival of vascular plants (Koukol and Conn, 1961; Jones, 1984; Lewis and Davin, 1999). Lignins, phytoalexins, flavonols, anthocyanins, and proanthocyanidins are all biosynthesized via branch
pathways of the phenylpropanoid pathway and are related to morphology, lodging resistance, disease resistance, pollinator attracting ability, and seed (coat) traits. Phenylalanine ammonia-lyase (PAL; EC 4.3.1.5) catalyzes the initial step in the biosynthesis of phenylpropanoids. There are 4 PAL homologs in the Arabidopsis genome. Transcription analyses and protein expression studies have shown that AtPAL1 and AtPAL2 are more important, and that AtPAL4 has limited importance, whereas AtPAL3 has little function (Wanner et al., 1995; Cochrane et al., 2004).

Brassica is a remunerative genus within the family Brassicaceae/Cruciferae. Many members of this genus are grown widely as vegetable crops, including cabbage (B. oleracea var. capitata), Chinese cabbage (B. rapa L. subsp. pekinensis), tuber mustard (B. juncea var. tumida), and turnip (B. rapa var. rapa), and as oil crops, including rapeseed (B. napus), B. rapa syn. campestris var. oleifera, and B. juncea. Rapeseed represents one of the world's major sources of vegetable oil. Molecular breeding of many desirable traits, including the yellow seed trait for good quality, strengthened and optimized lignin content for lodging and disease resistance, and enhanced phytoalexin content for repelling phytophagous pests, is expected in rapeseed genetic improvement. Research on B. napus PAL genes will promote investigation of the relationship of PAL genes to some important rapeseed traits at the molecular level and open the way to modification of PAL-related rapeseed traits through genetic engineering. It has been demonstrated that content and properties of lignin and its associated phenolics can be altered by overexpression or downregulation of enzymes involved in lignin biosynthesis (Boudet, 1998).

Yellow-seeded B. napus has a thinner seed coat and lower hull percentage, which result in higher seed oil and meal protein content, and lower crude fiber content than black-seeded counterparts with the same genetic background (Abraham and Bhatia, 1986). Development of yellow-seeded rapeseed varieties is vigorously pursued in rapeseed quality improvement. To date, how the yellow seed trait of B. napus is formed is not well understood; however, yellow- and black-seeded rapeseed lines show differences in the catalytic level of common phenylpropanoid pathway enzymes (Ye et al., 2001) and PAL activity is drastically suppressed in yellow-seeded rapeseed lines, as compared with black-seeded lines (Liang and Li, 2004). We previously reported the isolation and characterization of 2 B. napus PAL1 full-length cDNAs, and that their transcription is down-regulated in late-stage seeds, possibly due to pathway feedback inhibition (Ni et al., 2008). Herein we report the cloning and functional characterization of 2 B. napus PAL2 genes, providing more evidence of their evolutionary and functional features, and their involvement in yellow seed trait formation.

Materials and methods

Plant materials and nucleic acid isolation

Seeds of B. napus stock lines 5B, L1, and L2 were maintained by the Chongqing Rapeseed Engineering Research Center, Beibei, Chongqing China, planted under normal experimental conditions, and selfed by anthotaxy-bagging. With typical agronomical traits, black-seeded 5B is bred as a maintainer of a male sterile line. Black-seeded L1 and yellow-seeded L2 are seed color near-isogenic lines. 5B, L1, and L2 have all been successively selfed for more than 11 generations. The parent seeds of each line in this study were collected from a selfed single plant. 5B roots, hypocotyls, cotyledons, stems, leaves, flower buds, flowers, and silique pericarps, and 10D (days after flowering), 20D, and 30D seeds were collected from representative plants (Ni et al., 2008). Only L1 and L2 reproductive organs, such as buds and flowers, and 10D, 20D, and 30D seeds were sampled from representative plants. Samples were immediately frozen in liquid nitrogen and stored at ~80 °C for isolating total RNA with a cetyltrimethylammonium bromide (CTAB)-based method (Zeng and Yang, 2002). Total genomic DNA was isolated from fresh leaves of representative plants of each line following the protocol of Tulsieram et al. (1992).

5' and 3' RACE of BnPAL2 genes

From the roots of 5B, 5 μg of total RNA was used for RACE handling, using a GeneRacer™ kit (Invitrogen, USA) according to the manufacturer's instructions, and the first-strand total cDNA was used as a template for amplifying 5' and 3' cDNA ends of BnPAL2 genes. For 5' RACE, antisense primer RPAL5-2 (5'-CTTTGTTCATACTCTCCTCATAACCC
AATCACT-3') corresponding to a 5' conservative site of AtPAL2 was paired with kit primer 5'P (5'-CGACTGGAGCACAGGACTGA-3'). For 3' RACE, sense primer FPAL3-2 (5'-TGTTCTCAACCAGATGTGGAGCTCC-3') corresponding to a 3' conservative site of AtPAL2 was paired with kit primer 3'P (5'-GCTGTAACGATACGCTACGTAACG-3'). The 50-μL PCR contained 2.5 units of Taq DNA polymerase and 1 μL of total first-strand cDNA. The PCR procedure was as follows: 2 min at 94 °C, 35 cycles of 94 °C for 1 min, 58 °C (3' RACE) or 55 °C (5' RACE) for 1 min, and 72 °C for 1 min, with a final incubation of 10 min at 72 °C for full dA-tailing of the PCR products. After agarose gel electrophoresis and ethidium bromide staining, the specific bands were recovered and ligated with the pMD18-T vector (TaKaRa, Dalian, China). After CaCl2 transformation (Sambrook et al., 1989), DH5α colonies were checked by PCR and representative clones of each insert length were sequenced using primers M13F/M13R by BioAsia Biotechnology, Shanghai, China.

Amplification and analyses of full-length cDNAs and genomic sequences of BnPAL2-1 and BnPAL2-2

Based on the sequencing results of 5' RACE, upstream gene-specific primers FPAL2-1 (5'-AGGAACACAAAACCAAACTTC-3') and FPAL2-2 (5'-AGGAACACAAACTTCCAAATCTC-3') were designed as the utmost 5' ends of 2 clones, signifying 2 unique genes. Based on the sequencing results of 3' RACE, downstream gene-specific primers RPAL2-1 (5'-GAGAATTACACGTAATTAC-3') and RPAL2-2 (5'-AGTTACGTCGTAATCATAC-3') were designed as the utmost 3' ends of 2 clones, signifying 2 unique genes. FPAL2-1 and FPAL2-2 were combined with RPAL2-1 and RPAL2-2, respectively, to form 4 primer pairs for amplifying full-length cDNAs of BnPAL2 genes using 50 μL of standard Taq-PCR containing 1 μL of total first-strand cDNA as a template. PCR cycling conditions were as follows: 2 min at 94 °C, 35 cycles of 94 °C for 1 min, 53 °C for 1 min and 72 °C for 3 min, followed by 10 min at 72 °C. The same conditions were used to amplify genomic sequences of BnPAL2-1 and BnPAL2-2 by substituting the template with 1 μL of total genomic DNA of 5B, L1, and L2.

Electrophoresis detection, gel recovery, T-vector cloning, and sequencing were conducted.

Vector NTI Advance v.9.0 was used for ORF translation, sequence alignment, and calculation of protein and nucleotide parameters. ClustalX v.2.0, Phylip-3.68, and TreeView were used for construction and editing of the phylogenetic tree based on protein sequences. Other bioinformatic analyses were performed online through website links provided by http://www.ncbi.nlm.nih.gov and http://www.expasy.org/tools/.

Southern blot analysis

A 722-bp coding region conserved in both BnPAL2-1 and BnPAL2-2 was amplified using BnPAL2-2 full-length cDNA as a template, and primer pair FBNP2A (5'-AGAGCTCAACACGAAGGAGACATGCCACAC-3')/RBNP2A (5'-TCTAGATCGTTAATGGAACGGAGCTCC-3'). The amplified fragment was recovered and labeled by random-primed incorporation of Digoxigenin-labeled deoxyuridine-triphosphate (DIG-dUTP), using a DIG DNA Labeling and Detection Kit (Roche, Germany). Genomic DNA of 5B (40 μg) was fully digested with enzymes DraI, EcoRI, and EcoRV in separate reactions, then fractionated by electrophoresis on a 0.8% agarose gel and transferred onto a positively charged nylon membrane (Roche) through capillarity (Sambrook et al., 1989). DNA was fixed on the membrane by baking at 80 °C for 2 h under vacuum. Hybridization was performed in a DIG Easy Hyb (Roche) oven. Moderate-stringency hybridization was performed at 44 °C for 16 h, followed by washing in 0.5 × SSC and 0.1% SDS. High stringency hybridization was performed at 54 °C for 16 h, followed by washing in 0.1 × SSC and 0.1% SDS. Immunological detection was performed using Anti-Digoxigenin-AP and NBT/BCIP (DIG DNA Labeling and Detection Kit, Roche). The wet filters containing colored bands were photographed.

Analysis of gene expression by semi-quantitative RT-PCR

For analysis of BnPAL2-1 and BnPAL2-2 expression patterns in different organs of B. napus, 2.5 μg of total RNA from each organ of 5B was reverse-transcribed into first-strand total cDNA (10 μL in volume) using an RNA PCR Kit (AMV) v.3.0 with
Oligo dT-Adaptor Primer (TaKaRa). Gene-specific sense primers FBNP2-1E (5'-AGCATGCATG GATCATCAACGGATCGTT-3') and FBNP2-2E (5'-AGCATGCATGGATCAGACTAACGGTT AGTTCC-3') were paired with antisense primer RBNP2A to amplify 1172-bp and 1169-bp cDNA fragments of BnPAL2-1 and BnPAL2-2, respectively. One microliter of the total first-strand cDNA was used as a template in a 50 μL standard Taq PCR. The amplification steps were as follows: 2 min at 94 °C, 27 cycles of 1 min at 94 °C, 1 min at 62 °C, and 90 s at 72 °C, followed by 10 min at 72 °C. For internal control, primers F26S (5'-CACAATGATAGGAAGAGCCGAC-3') and R26S (5'-CAAGGGAACGGGCTTGGCAGAATC-3') were designed to amplify the house-keeping gene orthologous to A. thaliana 26S rRNA (Singh et al., 2004). To compare transcript abundance of BnPAL2-1 and BnPAL2-2 between L1 and L2, total RNA samples from buds, flowers, 10 D seeds, 20 D seeds, and 30 D seeds of both lines were reverse-transcribed and amplified for 27 cycles using primer pairs FBNP2-1E/RBNP2A and FBNP2-2E/RBNP2A, respectively, with other conditions the same as described above. Agarose gel electrophoresis detection was performed; each reaction was conducted twice.

**Prokaryotic expression of BnP AL2-1 and BnP AL2-2**

Using corresponding full-length cDNA clones as templates, primers FBNP2-1E and RBNP2-1E were used to amplify the ORF of BnP AL2-1, while FBNP2-2E (5'-AGCATGCATG GATCATCAACGGTT AGTTCC-3') and RBNP2-2E (5'-AGGTACCTTAA CATCCGGAAT CGGAGCTCC-3') were used to amplify the ORF of BnP AL2-2. The PCR products were recovered, T-vector cloned, and sequenced. These 2 fragments were subcloned into vector pQE30 (The QIAexpressionist™, Qiaogen, Germany) using Sphi/KpnI sites. *E. coli* strain M15 was transformed with recombinant plasmids. For each gene, 1 positive M15 clone was cultured in 20 mL of Luria-Bertani (LB) broth supplemented with ampicillin (100 μg mL⁻¹) and kanamycin (25 μg mL⁻¹) for 12 h at 37 °C. Bacteria were grown in 300 mL of LB with antibiotics at 37 °C until reaching an OD₆₀₀ of 0.6, when a final concentration of 0.5 mM isopropyl thio-β-D-galactoside (IPTG) was added. Bacteria were then cultured at 22 °C for 16 h. Cells were harvested by centrifugation at 4000 x g for 20 min and stored at -80 °C until required. Purification of 6x His-tagged recombinant proteins was performed using Qiagen Ni-NTA columns, following the QIAexpressionist™ manual.

Protein concentrations were determined by the Bradford (1976) method, using BSA as the standard. BnP AL2-1 and BnP AL2-2 proteins purified from the bacterial homogenate were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), using the mid-range protein marker BM523 (BBI, Canada), and were visualized by staining with Coomassie Blue. The activity was assayed spectrophotometrically according to the method described by Rosler et al. (1997). The reaction mixture containing 2 mL of 40 mM borate buffer (pH 8.7), 1 mL of 20 mM L-phenylalanine, and 0.15 mL of purified enzyme was incubated for 30 min at 37 °C, and the reaction was terminated with 0.85 mL of 1 M HCl. The reaction mixture was centrifuged at 2000 x g for 5 min before reading at 290 nm. One unit of enzyme was defined as the formation of 1 μg of trans-cinnamic acid 1 h⁻¹ at 37 °C under the conditions of assay. The specific activity was expressed as units mg⁻¹ of protein.

**Results**

Isolation and molecular characterization of BnP AL2-1 and BnP AL2-2 genes

5' RACE PCR amplification resulted in a wide band around 500 bp with some smear, while 3' RACE PCR generated a wide band around 250 bp. The 2 bands were gel-recovered and T-vector cloned, and with PCR screening the colonies of both bands showed insert length polymorphism. In total, 5 and 4 colonies of representative insert lengths of 5' and 3' RACES, respectively, were sent for sequencing. After excision of the GeneRacer RNA Oligo linker sequence, the net lengths of the five 5' RACE clones were 377-492 bp. NCBI BLASTn (Altschul et al., 1997) indicated that 2 of them showed no homology to PAL genes, while the other 3 were homologous to AtPAL2 mRNA (NM_115186) and defined 5' cDNA ends (448 bp and 453 bp) of 2 unique genes. After excision of
the 3’P/3’NP sequence, the net lengths of the 3’RACE clones were 110-374 bp. NCBI BLASTn indicated that 2 of them were not homologous to PAL genes, while the other 2 were homologous to AtPAL2 mRNA (NM_115186) and also defined 3’ cDNA ends (197 bp with 22-bp poly(A) and 194 bp with 25-bp poly(A)) of 2 unique genes.

Based on 5’ and 3’ RACE sequencing results, we designed 2 sense primers to pair with 2 antisense primers to carry out end-to-end amplification of full-length cDNAs of BnPAL2 genes using root total cDNA of 5B as a template. Electrophoresis indicated that primer pairs FPAL2-1/RPAL2-1 and FPAL2-2/RPAL2-2 both amplified a specific band of about 2.4 kb, while primer pairs FPAL2-1/RPAL2-2 and FPAL2-2/RPAL2-1 did not yield any band. By substituting the template with total genomic DNA of 5B, primer pairs FPAL2-1/RPAL2-1 and FPAL2-2/RPAL2-2 both successfully amplified a specific band of about 2.7 kb. After gel-recovery and T-vector cloning, sequencing results of 2 colonies were identical to each other for the 2 bands. Here, the gene from primer pair FPAL2-1/RPAL2-1 is denoted as BnPAL2-1, and from primer pair FPAL2-2/RPAL2-2 it is denoted as BnPAL2-2.

The sequenced full-length cDNAs of BnPAL2-1 and BnPAL2-2 were 2422 bp and 2421 bp, with corresponding genomic sequences of 2735 bp and 2742 bp, respectively. Figure 1 shows the gene structures and deduced amino acid sequences of BnPAL2-1 and BnPAL2-2. They both have a single intron—313 bp in BnPAL2-1 (540-852) and 321 bp in BnPAL2-2 (545-865). The introns follow standard GT…AG splicing borders, and their positions correspond to that of AtPAL2 (L33678), all interrupting a codon for a conservative Arg (R135 of AtPAL2, R142 of BnPAL2-1, and R141 of BnPAL2-2). BnPAL2-1 mRNA has a 114-bp 5’ untranslated region (UTR), a 2175-bp open reading frame (ORF), and a 133-bp 3’ UTR (poly(A) tail not included). BnPAL2-2 mRNA has a 122-bp 5’ UTR, a 2172-bp ORF, and a 127-bp 3’ UTR. No canonical polyadenylation signal AATAAA exists in the 3’ UTRs of BnPAL2-1 and BnPAL2-2. Both BnPAL2-1 and BnPAL2-2 share the typical feature of most functional genes, i.e. higher A+T contents in non-coding regions than in coding regions. The A+T contents in 5’ UTR, the intron, and the 3’ UTR of BnPAL2-1/BnPAL2-2 are 64.0%/64.0%, 68.0%/69.4%, and 66.2%/70.9%, versus 46.6%/45.3% in exon 1 and 48.6%/48.5% in exon 2, respectively.

When aligned on full-length genomic and cDNA scales, BnPAL2-1 shows 87.0% and 89.6% of identity with BnPAL2-2, respectively. Their 5’ UTR, exon 1, intron, exon 2, and 3’ UTR show 76.2%, 89.4%, 69.5%, 92.1%, and 70.1% of identity to each other, respectively. NCBI BLASTn indicates that BnPAL2-1 and BnPAL2-2 are homologous (local identities in parentheses) to AtPAL2 (1711/2022 = 84% and 1648/1890 = 87%), AtPAL1, and BnPAL1 genes (1467/1748 = 83% to 1511/1789 = 84%), and AtPAL4 (1282/1755 = 73% and 1297/1758 = 73%), AtPAL3 (1225/1770 = 69% and 1232/1764 = 69%), and PAL genes from non-cruciferous species (less than 1307/1748 = 74%). Calculated on Vector NTI, BnPAL2-1 is 79.0% identical to AtPAL2, with identities in 5’ UTR, exon 1, intron, exon 2, and 3’ UTR of 53.8%, 79.3%, 49.0%, 86.9%, and 72.5%, respectively. BnPAL2-2 is 78.9% identical to AtPAL2, with identities in 5’ UTR, exon 1, intron, exon 2, and 3’ UTR of 53.8%, 79.3%, 49.0%, 86.9%, and 72.5%, respectively. BnPAL2-1 and BnPAL2-2 are more conserved in coding regions than in non-coding regions, and exon 2 is more conserved than exon 1.

**Analysis of the deduced BnPAL2-1 and BnPAL2-2 proteins**

Deduced BnPAL2-1 polypeptide consists of 724 amino acid (aa) residues, with a molecular weight (Mw) of 78.52 kD and an isoelectric point (pI) of 6.03, while the 723-aa BnPAL2-2 has a Mw of 78.45 kD and a pI of 5.97. They share 96.4% identity and 97.8% similarity with each other. NCBI BLASTp indicated that they show greater homology to AtPAL2 (NP_190894) than to AtPAL1 (NP_181241), BnPAL1-1 (ABC69916), BnPAL1-2 (ABC69917), AtPAL4 (AAP59440), and AtPAL3 (NP_196043). In full-length pairwise alignment, BnPAL2-1 shows 92.7%/95.2%, 89.9%/93.5%, 89.6%/93.8%, 89.9%/93.4%, 78.1%/86.1%, and 69.6%/78.4% of identity/positive to AtPAL2, AtPAL1, AtPAL4, and AtPAL3, respectively; for BnPAL2-2 the values are 92.7%/94.9%, 90.4%/93.9%, 90.2%/93.9%, 90.3%/93.9%, 77.9%/85.9%, and 69.1%/78.4%, respectively. BnPAL2-1 and BnPAL2-2 are more homologous to AtPAL2 than to other AtPAL proteins and paralogs.
Isolation and functional characterization of two rapeseed genes orthologous to Arabidopsis thaliana Phenylalanine Ammonia-lyase 2

**BnPAL2 - 1**

AGGACGACATTGCAAAACTGATCATCGAACGAGTATGCCATGGTTGAGGATGCTGAAAGCTATGACGGTTTCGGAGCCGTTTGCTCTCTGGGAGAGGGTTTGATAAGGTGTTCAC

**BnPAL2 - 2**

AGGACGACATTGCAAAACTGATCATCGAACGAGTATGCCATGGTTGAGGATGCTGAAAGCTATGACGGTTTCGGAGCCGTTTGCTCTCTGGGAGAGGGTTTGATAAGGTGTTCAC

Figure 1. Nucleotide and deduced amino acid sequences of BnPAL2-1 (A) and BnPAL2-2 (B). The start codon ATG and the stop codon TAA are in bold face and are solid-underlined. The intron is denoted with a gray background. The underlined last nucleotide represents the polyadenylation site. The predicted PAL conserved domain V₁₇-N₁₂₀ of BnPAL2-1 or V₁₇-V₁₉₉ of BnPAL2-2 is underlined. Residues corresponding to active sites in *Petroselinum crispum* PAL1 (PsPAL1, CAA68938) and *Rhodosporidum toruloides* PAL (RtPAL, X51513) are boxed, and those of the catalytic center are further denoted with a gray background.
BnPAL1-1/BnPAL1-2. They also show certain homologies to PALs from other plant families (identities less than 604/724 = 83%, positives less than 659/724 = 91%), fungal PALs (identities less than 257/721 = 35%, positives less than 385/721 = 53%), and histidine ammonia-lyase (HAL) from microbes and animals (identities less than 170/476 = 35%, positives less than 258/476 = 54%) (Figure 2). Based on phylogenetic analysis, BnPAL2-1 and BnPAL2-2 are clustered with AtPAL2 first, and BnPAL1-1 and BnPAL1-2 are clustered with AtPAL1 first. These 2 mini-groups are further clustered with other dicot PALs, AtPAL4, and AtPAL3 to form a dicot PAL group. PALs from monocots are near the dicot PAL group. The gymnosperm-fern PAL group, the fungi PAL group, and the animal-microbe HALs show increasing distances from the angiosperm PALs (Figure 2). In multi-alignment, BnPAL2-1 and BnPAL2-2 are more similar to AtPAL2 than to AtPAL1, especially at many specific residues, e.g. the underlined I/E, N/V, E/A, A/E, N/G, A/S, N/K, V/T, K/T, S/N, S/C, I/V, F/Y, N/H, T/S, T/A, L/Q, I/V, V/I, R/A, Y/F, A/V, I/M, V/S, S/T, H/Q, D/G, S/A, E/K, T/V, I/M, M/L, and N/K in Figure 3. It is clear that BnPAL2-1 and BnPAL2-2 are more orthologous to AtPAL2 than to AtPAL1.

An NCBI Conserved Domain (CD) (Marchler-Bauer and Bryant, 2004) search detected a conserved domain pfam00221 (PAL) located from V63 to N561 in BnPAL2-1 and from V62 to V559 in BnPAL2-2. All highly conserved residues, especially those within the catalytic center of plant PALs, are unchanged in BnPAL2-1 and BnPAL2-2, e.g. the residues corresponding to A202, S203, G204 that autocatalytically form the cofactor 4-methylidene-imidazole-5-one (MIO), N384, Q483, R550, N550, E600, L1381, L2060, Y351, N487, Q488, E484, Y110 (Schuster and Rétey, 1994; Röther et al., 2002; Ritter and Schulz, 2004) and R274 (Langer et al., 1997) from P. crispum PAL (Figure 3).

SignalP v.3.0 did not predict any signal peptide in BnPAL2-1 or BnPAL2-2. WoLFPsort (http://wolfpsort.seq.cbrc.jp/) predicted that these proteins were located in the cytoplasm, conforming to the results of previous studies (Jones, 1984; Shaw et al., 1990). TMpred (Hofmann and Stoffel, 1993) predicted, however, that BnPAL2-1 has 2 strong transmembrane helices at regions T71-E91 and V266-E282, and BnPAL2-2 has 4 at regions N65-G57, A265-E283, E281-V301, and D314-D332. This raises the question of whether plant PALs have an association with the cellular membrane system, even though they are targeted to the cytoplasm. A NetPhos v.2.0 search (Blom et al., 1999) predicted 37 significant potential phosphorylation sites (16 for S, 12 for T, and 9 for Y) in both BnPAL2-1 and BnPAL2-2, suggesting that phosphorylation may be involved in functional regulation of the proteins. NetN Glyc v.1.0 predicted 3 significant potential N-glycosylation sites in BnPAL2-1 (N19, N197, and N268) and BnPAL2-2 (N5, N195, and N267), though proteins without signal peptides may not be glycosylated (in vivo). The fungal PAL from Ustilago maydis did not show any evidence of glycosylation (Kim et al., 1996), while the wound-induced PAL in potato showed absolute necessity for glycosylation (Shaw et al., 1990).

SOPMA (Geourjon and Deléage, 1995) predicted that BnPAL2-1 and BnPAL2-2 contained 50.83% and 50.21% alpha helices, 29.42% and 30.15% random coils, 11.33% and 11.07% extended strands, and 8.43% and 8.58% beta turns, respectively. Several obvious large helices exist in these 2 proteins, e.g. T509-K554 of BnPAL2-1 and T508-K554 of BnPAL2-2. The SwissModel (Arnold et al., 2006) predicted that BnPAL2-1 and BnPAL2-2 share nearly identical tertiary structures (Figure 4), both showing 83% of identity to model 1w27b.pdb, which is the subunit of the suggested homotetrameric quaternary structure of parsley PAL detected by X-Ray crystal diffraction at a resolution of 1.7 Å (Ritter and Schulz, 2004). BnPAL2-1 and BnPAL2-2 conform to typical features of PALs.

**In vitro catalytic activity of E. coli-expressed His-tagged BnPAL2-1 and BnPAL2-2**

The E. coli-expressed BnPAL2-1 and BnPAL2-2 proteins with N-terminal 6× His-tag were purified by Ni-NTA affinity chromatography. In SDS-PAGE they showed molecular weights that were in close agreement with the calculated values (80.09 kD for BnPAL2-1 and 80.02 kD for BnPAL2-2); however, as with BnPAL1-1 (Ni et al., 2008), SDS-PAGE analysis of BnPAL2-1 after successive purification steps still showed an extra band of higher Mw, while BnPAL2-2 showed only the predicted band (Figure 5). After the first, second, and third rounds of purification,
Figure 2. Phylogenetic tree of BnPAL2-1 and BnPAL2-2. Multi-alignment of protein sequences was made using ClustalX v.2.0, and the tree was drawn using Phylip-3.68 with the Neighbor-Joining (NJ) algorithm. Bootstrap values (100 replicates) are shown at branch points.
Figure 3. Multi-alignment of 4 BnPALs, 4 AtPALs, PcPAL1, and RtPAL. Identical, relatively conservative, slightly similar, and non-similar residues are denoted by reverse display, dark gray background, light gray background, and white background, respectively. Asterisks in the upper line mean that the corresponding residues in BnPAL2-1 and BnPAL2-2 are more similar to AtPAL2 than to AtPAL1. In the consensus line, the catalytic center or active site residues are underlined.
BnPAL2-1 had a protein yield of 0.305, 0.278, and 0.179 mg, and a tested activity of 2.680, 3.000, and 3.720 U mg$^{-1}$, while BnPAL2-2 had a protein yield of 0.129, 0.171, and 0.107 mg, and a tested activity of 25.286, 144.000, and 214.667 U mg$^{-1}$, respectively. This indicates that both BnPAL2-1 and BnPAL2-2 encode bioactive proteins catalytically competent for deamination of L-phenylalanine; however, under the same assay conditions, BnPAL2-1 showed much less activity than BnPAL2-2.

Possible BnPAL2 genes revealed by Southern hybridization

Using BnPAL2-1 cDNA fragment as a probe, Southern hybridization of B. napus total genomic DNA under moderate stringency conditions generated 5-6 bands in each restriction digestion, whereas high stringency Southern detection yielded 2 bands for each enzyme digestion (Figure 6). These results imply that B. napus probably has 2-6 BnPAL2 genes, and that BnPAL2-1 and a highly homologous
paralog (BnPAL2-2 or not) are closer to each other than they are to other BnPAL2 paralogs.

**Organ-specificity of BnPAL2-1 and BnPAL2-2**

Transcripts of BnPAL2-1 and BnPAL2-2 were detectable in all tested organs, but organ-specificity still existed in both genes (Figure 7). Transcription of BnPAL2-1 was most abundant in buds, with sequentially decreased abundance in stems, roots, flowers, seeds of all stages, silique pericarps, and leaves, and was weak in hypocotyls and cotyledons. Expression of BnPAL2-2 was most intensive in roots, declining sequentially in buds, stems, flowers, and 30D seeds, and was weak in all other organs. BnPAL2-2 is obviously more organ-specific than BnPAL2-1. Though they both are expressed considerably in roots, flower buds, stems, and flowers, and are weakly expressed in hypocotyls and cotyledons, BnPAL2-1 is dominant over BnPAL2-2 in seeds of all stages, silique pericarps, leaves, and even stems, whereas BnPAL2-2 is more abundant in roots.

**Status of BnPAL2-1 and BnPAL2-2 in yellow-seeded line L2**

Primer pairs FPAL2-1/RPAL2-1 and FPAL2-2/RPAL2-2 both successfully amplified a specific band of about 2.7 kb from near-isogenic black-seeded line L1 and yellow-seeded line L2, respectively. Sequencing results indicated that both L1 and L2 have complete structures of BnPAL2-1 and BnPAL2-2. L1 and L2 have identical sequences of the 2 genes, which diverge from those of 5B for a few nucleotides of nonsense variations; however, transcription levels of BnPAL2-1 and BnPAL2-2 in L2 buds, flowers, and seeds of all 3 stages were obviously lower than those of L1 (Figure 7).
Discussion

PAL is encoded by a multi-gene family in most plant species, e.g. there are 4 PAL genes in *A. thaliana* (Wanner et al., 1995; Cochrane et al., 2004) and potato possibly contains over 40 PAL genes (Joos et al., 1992). It is assumed that the genome of the *Brassica* progenitor has been triplicated after the *Arabidopsis-Brassica* split, followed by complicated rearrangements, including loss of some of the triplicated genes (Lagercrantz, 1998; Schmidt et al., 2001; Johnston et al., 2005; Town et al., 2006; Yang et al., 2006). It was anticipated that *B. rapa* and *B. oleracea* would both have about 3 genes corresponding to each gene from *A. thaliana* and 6 for the amphidiploid *B. napus*. Here, Southern blot analysis with high and moderate stringencies yielded at most 2 and 6 hybridization bands, respectively, suggesting that there are 2-6 BnPAL2 genes orthologous to *AtPAL1* genes. This is very similar to the BnPAL1-*AtPAL1* relationship (Ni et al., 2008), suggesting that both *PAL1* and *PAL2* loci were most probably triplicated in the *Brassica* progenitor. From the phylogenetic tree it can be seen that BnPAL1-1 and BnPAL1-2 are orthologs of AtPAL1, and BnPAL2-1 and BnPAL2-2 are orthologous to AtPAL2. The gene duplication event that generated the AtPAL3/AtPAL4 progenitor and AtPAL1/AtPAL2 progenitor occurred early—before the formation of family Brassicaceae, or even before the divergence of many dicot families after the dicot-monocot split (Figure 2). Moreover, in the Brassicaceae ancestor, the duplication to generate the AtPAL1-AtPAL2 split occurred before the *Arabidopsis-Brassica* split. This implies that most Brassicaceae species should contain PAL genes orthologous to all 4 AtPAL genes.

In *A. thaliana*, AtPAL1 and AtPAL2 expression patterns are similar for various tissues, and both are suggested to be involved in lignification (Ohl et al., 1990; Mauch-Mani and Slusarenko, 1996; Costa et al., 2003; Rookes and Cahill, 2003; Cochrane et al., 2004). BnPAL2 and BnPAL1 members primarily mimic the corresponding AtPAL orthologs, i.e. both are involved in lignification and flavonoid biosynthesis, but functional division has emerged. Both BnPAL1 members are nearly constitutive and are most abundantly expressed in flowers, implying that they participate in important phenylpropanoid activities in most organs and are the preferential paralogs for flower pigmentation (Ni et al., 2008). BnPAL2 members diverge from BnPAL1 members and from each other. In contrast to BnPAL1, BnPAL2 members are more organ-specific. Though the organ specificity of BnPAL2-1 is somewhat similar to that of BnPAL1 genes, it has only weak expression in hypocotyls and cotyledons. BnPAL2-2 is more organ specific than BnPAL2-1. It is dominant in roots, but withdraws from many other organs, including seed coats. *A. thaliana* EST and genomic sequence analyses showed that AtPAL2 was most likely a candidate for lignification, whereas AtPAL1 may have no direct involvement in lignification (Costa et al., 2003). In *A. thaliana* the flower is not the organ with the most abundant PAL transcripts (Wanner et al., 1995), and no ESTs for any of the AtPAL isoforms were detected in flower buds (Costa et al., 2003). Besides in roots, all 4 BnPAL genes are abundantly transcribed in flowers and buds, implying that they are all important for both lignification and flower pigmentation. *A. thaliana* is self-pollinated and has white flowers, but rapeseed is cross-pollinated and has golden bright flowers. Apparently, the bright color of rapeseed flowers is a strong driving force for intensive biosynthesis of flavonoid pigments. This intergenus difference is also supported by clues from the *F3’Hi*/*TT7* locus (Xu et al., 2007). RT-PCR results revealed that, except for BnPAL2-2, the 3 other BnPAL genes are important for seed coat pigmentation.

The transcription of BnPAL2 genes, together with previous BnPAL1-1 and BnPAL1-2 studied, show less abundance in developing seeds in yellow-seeded line L2 than in black-seeded line L1, indicating the correlation of downregulation of PAL transcription to the yellow-seed trait. It is strange that L2 shows even less BnPAL2 expression than L1 in non-seed organs such as buds and flowers. In fact, BnPAL1-1 expression is obviously higher in buds and flowers of L2 than of L1 (Ni et al., 2008); therefore, PAL1+PAL2 overall expression in buds and flowers of L2 should be similar to that in L1. In L2 all 4 BnPAL genes keep intact coding capacity, but show the same trend of downregulation in developing seeds, implying that a seed-specific signal affecting PAL expression should be responsible for the yellow seed trait of L2.
Considerable difference in catalytic activity was detected between members of both subfamilies of BnPAL1 (Ni et al., 2008) and BnPAL2; in particular, BnPAL1-1 and BnPAL2-1 showed a 90-kD band accompanying the main band and decreased catalytic activity. These suggest that these 2 proteins are possibly easier to be inactivated by association with E. coli proteins or inefficient folding.

B. napus is the most widely grown oilseed rape in the world and benefits from biotechnology. The PAL genes isolated herein will undoubtedly speed up further investigation of the relationship between PAL genes and some important traits of rapeseed at the molecular level, and therefore lay the base for molecular breeding of PAL-related traits, such as yellow-seed and adversity tolerance.

Acknowledgements

This work was supported by National Basic Research (“973”) Program of China (2006CB101604), the National High Technology Research and Development Program of China (2006AA10A113 and 2006AA100106), National Natural Science Foundation of China (NSFC, No. 30330400), and Major Program of Chongqing Municipal Natural Science Foundation (8446).

References


Isolation and functional characterization of two rapeseed genes orthologous to Arabidopsis thaliana Phenylalanine Ammonia-lyase 2


