

RSC Advances



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. This *Accepted Manuscript* will be replaced by the edited, formatted and paginated article as soon as this is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

1 **Title:**

2 **Transcriptome-wide identification and characterization of *Ornithogalum***
3 ***saundersiae* phenylalanine ammonia-lyase gene family**

4

5 **Authors and affiliation:**

6 Zhi-Biao Wang, Xi Chen, Wei Wang, Ke-Di Cheng, Jian-Qiang Kong*

7 Institute of Materia Medica, Chinese Academy of Medical Sciences & Peking Union

8 Medical College (State Key Laboratory of Bioactive Substance and Function of

9 Natural Medicines & Ministry of Health Key Laboratory of Biosynthesis of Natural

10 Products), Beijing, 100050, China

11

12 *Author to whom correspondence should be addressed;

13 E-Mail: jianqiangk@imm.ac.cn TEL: (86-10)-63165169

14

15

1 **Abstract**

2 OSW-1 is a promising antitumor glycoside present in the plant *Ornithogalum*
3 *saundersiae*. Biosynthesis of the *p*-methoxybenzoyl group on the disaccharide moiety
4 of OSW-1 is known to take place biochemically by phenylpropanoid biosynthetic
5 pathway, but molecular biological characterization of related genes has been
6 insufficient. Phenylalanine ammonia-lyase (PAL, EC 4.3.1.24), catalyzing the
7 deamination of L-phenylalanine yielding *trans*-cinnamic acid, plays a key role in
8 phenylpropanoid metabolism. Thus, the study on the characterization of the genes
9 involved in the OSW-1 biosynthetic pathway, in particular the well-documented genes
10 such as PAL, is essential to further understanding the biosynthesis of OSW-1. Here,
11 transcriptomic sequencing of *O. saundersiae* was performed to speed up of the
12 identification of a large number of related genes of OSW-1 biosynthesis. *De novo*
13 assembly of the transcriptome sequence provided 210,733 contigs and 104,180
14 unigenes, and four unigenes showing high similarities with PALs were retrieved. Two
15 full-length cDNAs encoding PALs (*OsaPAL2* and *OsaPAL62*) from *O. saundersiae*
16 were cloned using sequence information from these four unigenes. The PAL and
17 tyrosine ammonia-lyase (TAL) activities of recombinant OsaPAL proteins were
18 unambiguously determined by HPLC with UV and MS detection, as well as by NMR
19 spectroscopy. Subsequently, a series of site-directed mutants were generated with the
20 aim of improving enzyme activity and to investigate the importance of particular
21 residues in determining substrate selectivity. The results reveal that the Phe-to-His
22 mutants, OsaPAL2F134H and OsaPAL62F128H, exhibited higher TAL activity than

1 the corresponding wild types providing direct evidence that the Phe residue is
2 responsible for substrate specificity. Mutagenesis studies also demonstrated that the
3 Thr-to-Ser mutants, OsaPAL2T196S and OsaPAL62T194S, showed significantly
4 higher substrate affinity than wild-types. The Gly-to-Ala mutants, OsaPAL2G209A
5 and OsaPAL62G207A, showed higher PAL and TAL activities. These findings
6 provide the further insight into the genes responsible for OSW-1 biosynthesis and will
7 facilitate the future application of OsaPALs in synthetic biology.

8 **Keywords** OSW-1; Phenylalanine ammonia-lyase; *Ornithogalum saundersiae*;
9 synthetic biology

10

11

12

13

14

15

16

17

18

19

20

21

22

1 **1. Introduction**

2 Phenylalanine ammonia-lyase (PAL, EC 4.3.1.24), a member of the ammonia-
3 lyases superfamily, catalyzes the conversion of L-phenylalanine to *trans*-cinnamic
4 acid by a non-oxidative deamination. In recent years, much attention has been paid
5 to the presence of PALs in medicinal plants since they are involved in the biosynthesis
6 of the majority of phenolic containing secondary metabolites, many of which are
7 pharmacologically active.^{1,2}

8 OSW-1 (Fig. 1A), a natural saponin isolated from *Ornithogalum saundersiae*,
9 exhibits exceptionally potent antitumor activity both *in vitro* and *in vivo*.³⁻⁵ As such,
10 it is a promising lead compound for the development of novel antitumor drugs.
11 However, because of its low content in the plant and a long and laborious synthesis,
12 ⁶⁻¹⁰ little progress has been made in developing OSW-1 as a potential drug candidate
13 since its first discovery in 1992.¹¹ Accordingly, we have initiated a project to find an
14 alternative, preparative scale method to produce OSW-1 and to elucidate its
15 biosynthetic pathway and the enzymes involved.

16 OSW-1 is characterized by a disaccharide moiety attached to the C-16 position of
17 the steroid aglycone which contains a *p*-methoxybenzoyl (MBz) and an acetyl (Ac)
18 group.¹¹ According to previous structure-activity relationship (SAR) studies, the
19 disaccharide moiety is important for the cytotoxicity of OSW-1 and, in particular,
20 removal of the Ac and MBz groups decreases its activity by approximately 1000-fold.
21 ^{3,12,13} Many previous studies have indicated that the *p*-hydroxybenzoic acid group
22 derives either from L-Phe or L-Tyr (Fig. 1B) through the respective PAL-catalyzed

1 formation of *trans*-cinnamic acid or *p*-coumaric acid.¹⁴⁻¹⁶ Thus, it is likely that PALs
2 play an important role in this phenylpropanoid biosynthetic pathway to OSW-1.
3 However, to date, none of the PALs in *O. saundersiae* has been cloned and
4 characterized.

5 PAL was first purified from *Hordeum vulgare* in 1961¹⁷ but was later shown to be
6 widespread in plants,¹⁸⁻²⁰ fungi²¹⁻²³ and prokaryotes.²⁴⁻²⁶ In plants, PALs are fairly
7 ubiquitous, and are found in monocots,²⁷⁻³⁰ dicots,^{1,18,19,31} gymnosperms,³²⁻³⁵ ferns,
8³⁶ lycopods,³⁶ liverworts,³⁶ and algae.³⁷ There is, however, no report about PALs in
9 Asparagaceae species.

10 Here, a gene family containing two OsaPAL genes was isolated for the first time
11 from an Asparagaceae plant species, *O. saundersiae*. After successfully functional
12 characterization, a series of site-directed mutants were generated with the aim of
13 improving enzyme activity and investigating the importance of specific residues in
14 determining substrate selectivity. The *in vitro* assays indicated that we successfully
15 improved both the PAL and TAL deamination activities of OsaPALs by single amino
16 acid substitutions, a discovery that we believe will provide additional ways to
17 improve PAL activity of other such enzymes. We also maintain that a greater
18 understanding of how OsaPALs participate in the biosynthesis of OSW-1 will
19 facilitate future applications of OsaPALs in synthetic biology.

20

21 **2. Experimental methods**

22 **2.1. Substrates, chemicals and enzymes**

1 Materials (suppliers) were as follows: L-Phe and L-Tyr used in enzyme assays
2 (Sigma-Aldrich Co. Ltd., St. Louis, MO, USA); In-Fusion® HD Cloning Kit and
3 restriction enzymes (Takara Shuzo Co. Ltd., Kyoto, Japan); Super RT cDNA Kit used
4 to synthesize full-length cDNAs and KOD-Plus-Neo DNA polymerases (Toyobo Co.
5 Ltd., Osaka, Japan); RNeasy Plant Mini Kit used for RNA extraction (Qiagen,
6 Dusseldorf, Germany); Ni-Sepharose (Invitrogen, Carlsbad, CA, USA); Fast
7 Mutagenesis System kit used for site-directed mutagenesis (TransGen Biotech Co.
8 Ltd., Beijing, China). All other fine chemicals used in this study were analytically
9 pure.

10

11 **2.2. Strains and plasmids**

12 Prokaryotic expression vector pET-28a (+) used for heterologous expression was
13 purchased from Novagen, Madison, USA. Expression host strain *Transetta* (DE3) and
14 the cloning vector *pEASY*®-Blunt were purchased from TransGen Biotech Co. Ltd.,
15 Beijing, China. TG1, for use as a bacterial host for recombinant plasmid amplification,
16 was grown in Luria-Bertani medium (10 g l⁻¹ Bacto-Tryptone, 5 g l⁻¹ Bacto-yeast
17 extract, 10 g l⁻¹ NaCl) or induced in TB medium (12 g l⁻¹ Bacto-Tryptone, 24 g l⁻¹
18 Bacto-yeast extract, 4 ml l⁻¹ glycerol, 72 mM K₂HPO₄, 17 mM KH₂PO₄)
19 supplemented with appropriate antibiotics for selection.

20

21 **2.3. Plant materials**

22 *O. saundersiae* was grown under sterile conditions on 6,7-V medium ³⁸ at a

1 temperature of 22°C and under a 16 h light/8 h dark cycle. Sterile bulbs were
2 collected and used immediately for RNA isolation.

3

4 **2.4. Transcriptome sequencing and analysis**

5 RNA extraction and cDNA library construction were done as described in Kong *et*
6 *al.*³⁹ The resultant cDNA library was sequenced using Illumina HiSeq™ 2000. Short
7 nucleotide reads obtained via Illumina sequencing were assembled by the Trinity
8 software (<http://www.trinity-software.com>) to produce error-free, unique contiguous
9 sequences (contigs). These contigs were ligated to obtain non-redundant unigenes,
10 which could not be extended on either end. Unigene sequences were aligned by Blast
11 X to protein databases like NCBI nr, Swiss-Prot, KEGG and COG (e-value <
12 0.00001), and aligned by Blast N to nucleotide databases nt (e-value < 0.00001),
13 retrieving proteins with the highest sequence similarity with the given unigenes along
14 with their functional annotations.

15

16 **2.5. Generation of full-length OsaPALs cDNA and sequence analysis**

17 Since the assembled sequences were products of *de novo* assemblies, they were
18 considered prone to error. To confirm that a sequence represented a true gene product,
19 experimental verification was performed by designing gene-specific primers (Table 1)
20 for the *OsaPAL* full-length sequences and verifying the identity of amplified products
21 by sequencing.

22 Total RNA isolated from sterile bulb tissue of *O. saundersiae* using an RNeasy

1 Plant Mini Kit (Qiagen) was used as templates for reverse transcription using primer
2 oligo (dT)₂₀ primers and reverse transcriptase ReverTra Ace (TOYOBO) according to
3 the manufacturer's instructions. Amplification of *OsaPAL* cDNAs was performed by
4 a nested PCR method using KOD Plus Taq polymerase and gene-specific primers
5 (Table 1). The amplified full-length cDNAs, *OsaPAL2* and *OsaPAL62*, were each
6 inserted into the *pEASY*[®]-Blunt vector to generate pEASY-*OsaPAL2* and
7 pEASY-*OsaPAL62* respectively for sequencing.

8 *OsaPAL2* and *OsaPAL62* were analyzed using online bioinformatic tools from
9 NCBI and ExPASy. Open Reading Frame (ORF) finding was performed by the
10 on-line program (<http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi>). Various
11 physicochemical parameters of proteins were evaluated using the ProtParam tool
12 (<http://web.expasy.org/protparam/>). TMHMM
13 (<http://www.cbs.dtu.dk/services/TMHMM/>) was used to predict the transmembrane
14 helices of proteins and SignalP 4.1 (<http://www.cbs.dtu.dk/services/SignalP/>) was
15 used to predict cleavage sites of signal peptides. Protein subcellular locations were
16 predicted using TargetP 1.1 Server (<http://www.cbs.dtu.dk/services/TargetP/>) and
17 functional sites of proteins were analyzed with PROSITE tools
18 (<http://prosite.expasy.org/>). Concord (<http://helios.princeton.edu/CONCORD/>)⁴⁰ was
19 used to predict the secondary structures of *OsaPAL2* and *OsaPAL62* and protein
20 multiple sequence alignment was performed using ClustalX (version 2.1).⁴¹ A
21 phylogenetic tree was constructed using the neighbor-joining method in the
22 MEGA5.1 program.⁴²

1 2.6. Protein expression

2 The prokaryotic expression vector pET-28a (+) was digested with restriction
3 endonucleases *Bam*HI and *Eco*RI to generate the linearized vector. Plasmids
4 pEASY-OsaPAL2 and pEASY-OsaPAL62 were used as templates for sub-cloning the
5 full-length sequences of *OsaPALs* with the respective specific primer pairs (Table 1).
6 After gel purification, the PCR product was sub-cloned into pET-28a (+) according to
7 the In-Fusion® HD Cloning Kit protocol. Finally, sequencing was used to verify the
8 integrity of plasmids pET28a–OsaPAL2 and pET28a–OsaPAL62, with both of them
9 containing a His₆ tag.

10 The expression plasmids pET28a–OsaPAL2 and pET28a–OsaPAL62 were
11 transformed into the expression host strain *Transetta* (DE3) grown in LB agar
12 containing 170 µg/ml chloromycetin and 50 µg/ml kanamycin. After overnight
13 incubation at 37°C, a single clone of each was used to inoculate 20 ml of
14 LB/chloromycetin/kanamycin at 37 °C, shaken at 200 rpm until the OD at 600 nm
15 reached a value of 1.0, diluted 1:50 into 50 ml TB supplemented with 170 µg/ml
16 chloromycetin and 50 µg/ml kanamycin, and shaken at 200 rpm at 37°C until the OD
17 at 600 nm reached 0.8. IPTG to 0.1 mM was added and, after shaking at 150 rpm at
18 20°C overnight, the induced cells were harvested by centrifugation (7,500 g, 2 min) at
19 4 °C. The pellets were then either stored at -80°C or used directly.

20

21 2.7. Enzyme purification

22 For enzyme purification, all steps were performed at 4°C. First, *E. coli* cells were

1 washed and re-suspended in lysis buffer (pH 8.0, 20 mM sodium phosphate
2 containing 5 mM imidazole and 300 mM NaCl). Cells were then lysed with a
3 high-pressure homogenizer (800 bar, 3 passes) after which 1 U/ml DNaseI was added
4 and the homogenate incubated at 4°C for approximately 2 h. After centrifugation at
5 10,000 g for 15 min, the supernatant was passed through a 0.2 µm pore-size filter to
6 remove *E. coli* cell debris and other contaminants and then loaded onto a
7 pre-equilibrated column containing Ni-NTA resin. The column was washed with
8 washing buffers (pH 8.0, 20 mM sodium phosphate buffer containing 20-50 mM
9 imidazole and 300 mM NaCl) to remove non-specifically bound proteins after which
10 elution buffer (pH 8.0, 20 mM sodium phosphate containing 300 mM imidazole and
11 300 mM NaCl) was used to elute the His₆-tagged protein.

12 To remove small molecules including imidazole, dialysis was performed. A 30 K
13 semipermeable membrane was selected and approximately 20 ml protein sample was
14 dialyzed against 1 litre dialysis buffer (pH 8.0, 10 mM sodium phosphate) for 4 h at
15 4°C with four changes of dialysis buffer. Proteins were then dried in a vacuum freeze
16 dryer and stored at -80°C until use.

17

18 **2.8. Enzyme activity and analysis**

19 Enzyme activity of recombinant purified OsaPALs was unambiguously determined
20 by a combination of HPLC-UV, HPLC-MS and ¹H and ¹³C NMR spectroscopies
21 using L-Phe and L-Tyr as substrates. The reaction mixture (200 µl) containing 0.1 M
22 CHES buffer (pH 9.5), 100 mM L-Phe or 10 mM L-Tyr and different amounts of

1 purified protein was incubated at 37°C for 30 min and terminated by adding 200 μ l
2 chloroform. After centrifugation at 15,000 g for 10 min, the supernatant was filtered
3 through an 0.2 μ m pore-size filter and analyzed by HPLC–UV on a HITACHI
4 LaCrom elite L-2000 HPLC system (HITACHI, Toyokawa, Japan) using a C18
5 column [YMC-Pack ODS-A (5 μ m, 12 nm, 250 \times 4.6 mm)] and gradient elution using
6 0.05% aqueous trifluoroacetic acid as solvent A and CH₃CN as solvent B. After
7 pre-equilibrated in 98:2 A: B (v/v), the sample was injected and chromatographed
8 using a linear gradient to 35:65 A: B (v/v) over 26 min at a flow rate of 1 ml/min

9 The reaction products of L-Phe (or L-Tyr) catalyzed by OsaPAL2 and OsaPAL62
10 [named OsaPAL2-L-Phe-p (or OsaPAL2-L-Tyr-p) and OsaPAL62-L-Phe-p (or
11 OsaPAL62-L-Tyr-p), respectively] were isolated using a YMC semi-preparative
12 column [YMC-Pack ODS-A (5 μ m, 12 nm, 250 \times 10 mm)] using 15 min linear
13 gradients of 40 to 65% solvent B for L-Tyr products and 40 to 80% solvent B for
14 L-Phe products, both at a flow rate of 2 ml/min. UV detection at 275 and 310 nm was
15 used for enzymatic products of L-Phe and L-Tyr, respectively. HPLC-MS was
16 performed using an Agilent 1200 RRLC series HPLC system (Agilent Technologies,
17 Waldbronn, Germany) coupled to a QTRAP tandem mass spectrometer (QTRAP
18 2000, Applied Biosystems/MDS SCIEX) equipped with a Turbo Ion spray source
19 (Concord, ON, Canada) operated in the negative ionization mode and controlled by
20 Analyst 1.5 software. Mass spectra were collected in the enhanced full mass scan
21 mode in the range m/z 100-1000. NMR spectra of products dissolved in deuterated
22 dimethyl sulfoxide (DMSO-d⁶) and placed in 5 mm NMR tubes were recorded using

1 Bruker AVIII-600 and Bruker AVIII-500 NMR spectrometers (Bruker-Biospin,
2 Germany) operating at 600 and 500 MHz respectively. Chemical shifts (δ) and
3 coupling constants (J) are given in ppm and hertz (Hz) respectively.

4

5 **2.9. Optimum pH of recombinant OsaPALs**

6 To evaluate the effect of pH on enzyme activity, the following buffers were used:
7 0.1 M MES, pH 5.5–6.7; 0.1 M HEPES, pH 6.8–8.2; 0.1 M CHES pH 8.6–10.0; 0.1
8 M CAPS, pH 9.7–11.1; 0.1 M Na₂HPO₄–NaOH, pH 11.5–12.0. Assays were
9 performed at a constant temperature of 37°C for 20 min and monitored continuously
10 with a multimode reader. Controls without enzymes were included and each
11 experiment was performed in triplicate.

12

13 **2.10. Optimum temperature of recombinant OsaPALs**

14 For determination of the optimal temperature, reactions were performed in 0.1 M
15 CHES buffer (pH 9.5) and pre-incubated at different temperatures in the range 30-60
16 °C for 10 min. Substrates were added to initiate the reactions and, after incubation for
17 15 min during which the change in absorption was measured, glacial acetic acid was
18 added to terminate them. Controls without enzymes were included and each
19 experiment was performed in triplicate.

20

21 **2.11. Kinetic analysis of recombinant OsaPALs**

22 Kinetic analysis of native enzymes was performed at pH 9.5 and 44°C in a total

1 volume of 200 μ l containing various concentration ranges of substrates (L-Phe 9.76
2 μ M to 20 mM; L-Tyr 15.6 μ M to 8 mM) and purified proteins (For L-Phe; 6.3 μ g
3 OsaPAL2, 15.5 μ g OsaPAL62, 47.32 μ g OsaPAL2F134H, 36.5 μ g OsaPAL62F128H,
4 19.6 μ g OsaPAL2T196S, 22.8 μ g OsaPAL62T194S, 23 μ g OsaPAL2V202S, 647.2 μ g
5 OsaPAL62V200S, 8.8 μ g OsaPAL2G209A, 12.1 μ g OsaPAL62G207A: For L-Tyr;
6 135.6 μ g OsaPAL2, 136.4 μ g OsaPAL62, 236.6 μ g OsaPAL2F134H, 182.84 μ g
7 OsaPAL62F128H, 175.2 μ g OsaPAL2G209A, 120.6 μ g OsaPAL62G207A). Assays
8 for individual substrates were performed in triplicate for 10 min (L-Phe) or 20 min
9 (L-Tyr). The formation of product was continuously monitored with the multimode
10 reader. Kinetic constants values were determined from Lineweaver–Burk plots. All
11 kinetic assays were performed in triplicate and controls without enzymes or substrates
12 were included.

13

14 **2.12. Site-directed mutagenesis**

15 Site-directed mutagenesis was used to investigate the importance of various amino
16 acid residues in determining enzyme activity. Point and double mutations were
17 introduced into OsaPALs by PCR-based amplification of the entire OsaPAL
18 expression plasmid (pET28a-OsaPAL2 or pET28a-OsaPAL62) using two mutated
19 oligonucleotide primers (Table 2), each complementary to the opposite strand of the
20 vectors. All components necessary for PCR-based mutagenesis were provided in the
21 Fast Mutagenesis System kit and were used according to the manufacturer's
22 instructions. All mutants were confirmed by sequencing and those plasmids with

1 target substitutions and without other unwanted mutations were retained. Triple
2 mutations were introduced in the same way using already existing mutants as
3 templates.

4

5 **2.13. Comparison of mutant activities**

6 To compare the activities of wild-type and mutants, mutants were expressed and
7 purified according to the methods described above. Purity of proteins was estimated
8 by SDS-PAGE and pure proteins were dried in a vacuum freeze-dryer and stored at
9 -80°C . Prior to use, enzymes were dissolved in CHES buffer (0.1 M, pH 9.5) and
10 protein determined using the Bio-Rad protein assay (Bio-Rad, USA). HPLC was used
11 to compare the catalytic activities of different mutants. Assays were performed in
12 1000 μl aqueous buffer containing 100 mM CHES (pH 9.5) and 25 μg enzyme. After
13 pre-incubation at 37°C for 10 min, substrates (50 mM L-Phe or 10 mM L-Tyr) were
14 added and reactions run for 1, 2 or 3 h before being terminated by boiling for 5 min.
15 Mixtures were then centrifuged (15,000 g, 10 min) and 15 μl injected into the HPLC
16 system. Each assay was performed in triplicate. Kinetic constants of some single
17 mutants were also determined using the above procedure.

18

19 **3. Results and Discussion**

20 **3.1. Transcriptome analysis of OsaPAL homology**

21 OSW-1 is a cholestane saponin, featuring a novel 3β , 16β , 17α
22 $-$ trihydroxycholest-5-en-22-one aglycone with an acylated disaccharide attached to

1 the 16-hydroxyl group (Fig. 1A). Biogenetic analysis showed there were at least six
2 kinds of enzymes responsible for OSW-1 biosynthesis, including terpenoid backbone
3 biosynthetic enzymes and steroid pathway enzymes resulting in OSW-1 aglycone
4 formation, cytochrome P450 hydroxylase able to add hydroxyl groups to the numbers
5 3, 16 and 17 of OSW-1 aglycone, glycosyltransferase involved in disaccharide moiety
6 attachment to 16-OH of OSW-1 aglycone, acyltransferases catalyzing introduction of
7 the acetyl and the 4-methoxybenzoyl groups on the disaccharide moiety, nucleotide
8 sugars biosynthetic pathway enzymes providing glycosyl donors in glycosylation
9 reactions and phenylpropanoid biosynthetic pathway enzymes converting aromatic
10 amino acids to 4-methoxybenzoyl group (Fig. 1A). A total of more than 40 enzymes
11 were deduced to involve in biosynthesis of OSW-1. It will take much more time to
12 isolate and further functionally characterize all of these genes by conventional
13 molecular biology technologies. Thus, it is particularly important to apply a
14 high-throughput method, allowing for drastically quicker and cheaper genes discovery,
15 and leading towards a far more comprehensive view of biosynthetic pathway of
16 OSW-1. The advent of next-generation sequencing approach such as transcriptomic
17 analysis provides a platform, which has been proved to be critical in speeding up of
18 the identification of a large number of related genes of secondary products. In the
19 previous investigation, about 40 contigs and unigenes were retrieved and annotated to
20 be responsible for phenylpropanoid biosynthetic pathway from transcriptome
21 sequence data of *O.saundersiae*.³⁹ Further batch alignment results revealed there are
22 four unigenes showing high similarities with PALs. These four unigenes, namely

1 25029, 25031, 26221 and 26880 were respectively 2150, 2130, 1409, and 317 bp in
2 length. A batch BlastX search of GenBank with the four unigenes indicated none of
3 these unigenes contained a full-length CDS. Of them, unigenes 26880 and 25029 had
4 the equal sequence identity with the same PALs deposited in GenBank, suggesting
5 they were situated within the same candidate PAL. It is the same case with unigene
6 26221 and 25031. Further sequence alignment analysis postulated unigenes 26880
7 and 25029 were 5'- and 3'-end of one candidate PAL, while unigenes 26221 and
8 25031 were 5' and 3'-end of another candidate PAL, respectively. Given all that, a
9 gene family containing two PAL cDNAs was acquired by transcriptomic analysis.

10

11 **3.2. Cloning and analysis of the full-length cDNAs encoding OsaPALs**

12 Two full-length members of the *PAL* gene family were isolated from *O.*
13 *saundersiae* by nested PCR using gene-specific primers (Fig. 2A and Table 1).
14 Identity between the cDNA sequences and the results of transcriptome sequencing
15 verified by sequencing indicated the presence of a *bona fide PAL* gene family in
16 plants. Sequence information for the two cDNAs, designated *OsaPAL2* and
17 *OsaPAL62*, was deposited in the GenBank database (*OsaPAL2* accession number
18 KF741222; *OsaPAL62* accession number KF741223). *OsaPAL2* was derived from the
19 unigenes 26221 and 25031 and contained an ORF of 2136 bp. *OsaPAL62* overlapped
20 with unigenes 26880 and 25029 at the 5' and 3' ends, respectively and was 2130 bp in
21 length. Sequencing identification proved the two cDNAs to be *bona fide PALs* of *O.*
22 *saundersiae* consistent with the fact that PALs are encoded by gene families in plant

1 species such as *R. idaeus*,⁴³ *Arabidopsis*⁴⁴ and *Populus trichocarpa*.⁴⁵ Moreover, due
2 to a lack of genome sequence, we cannot determine if the two cDNAs were derived
3 from alternative splicing.

4 The proteins encoded by *OsaPAL2* and *OsaPAL62* were predicted by ProtParam
5 tool to be polypeptides containing respectively 711 and 709 amino acids with
6 molecular weights of 77103.9 and 76570.6 Da. The instability indices (II) indicated
7 that the two proteins were stable (II < 40) and their low grand average of
8 hydropathicity (GRAVY) values of -0.139 (*OsaPAL2*) and -0.112 (*OsaPAL62*)
9 indicate their hydrophobicity.⁴⁶ The theoretical pI values of 5.64 (*OsaPAL2*) and 6.15
10 (*OsaPAL62*) were < 7 revealing their acidic nature. The TMHMM tool indicated that
11 *OsaPAL2* or *OsaPAL62* are not transmembrane proteins, and SignalP 4.1 showed no
12 signal peptide or signal peptide cleavage sites indicating they are not secreted proteins.
13 These results are consistent with the predictions of TargetP 1.1 that the two proteins
14 are not located in chloroplasts, mitochondria or in secretory pathways.

15 According to the results provided by PROSITE tools, both *OsaPALs* contain a
16 17-amino acid motif (GTITASGDLVPLSYIAG) which is a conserved signature in the
17 superfamily of ammonia-lyases including PAL, tyrosine ammonia-lyase (TAL, EC
18 4.3.1.23) and histidine ammonia-lyase (HAL, EC 4.3.1.3). The motif contains an
19 active Ala-Ser-Gly (A-S-G) triad which can be autocatalytically converted into a
20 MIO (4-methylidene-imidazole-5-one) ring serving as a prosthetic group and
21 activating substrates by means of electrophilic interactions.⁴⁷⁻⁴⁹ With respect to
22 predicted secondary structure, both proteins had 8 β -sheets and 25 α -helices, with the

1 residues differing between them lying primarily in random coils (Fig.3).

2

3 **3.3. Multiple sequence alignment and phylogenetic analysis**

4 BlastP in NCBI (<http://www.ncbi.nlm.nih.gov/BLAST>) and multi-alignment
5 analysis according to CLUSTAL X algorithm indicate the deduced polypeptide of
6 OsaPAL2 is highly homologous to that of MbPAL (BAG70992.1) in *Musa balbisiana*.
7 OsaPAL62 is highly homologous to PALs of some species e.g., it shares 85% identity
8 with the PAL sequence of *Cinnamomum osmophloeum* CoPAL (AFG26322.1) while
9 the OsaPAL2 and OsaPAL62 sequences are 83.6% homologous.

10 Phylogenetic analysis revealed that the PALs of angiosperm plants are divided into
11 two clusters with those from dicotyledonous plants [e.g. PIPAL (AFI71896.1) from
12 *Paeonia lactiflora* and RcPAL (AGH13333.1) from *Rhus chinensis*] being in one and
13 those from monocotyledon plants [e.g. MbPAL (BAG70992.1) from *Musa balbisiana*]
14 in the other (Fig. 4). Amino acid sequences of OsaPAL2 and OsaPAL62 belong to the
15 latter cluster. It has been reported that PALs from monocotyledons always possess
16 TAL activity^{50,51} suggesting OsaPAL2 and OsaPAL62 may have PAL/TAL activities.
17 In the resulting phylogeny, OsaPAL62 and PALs from *Lycoris radiata* (LrPAL) and
18 *Dendrobium candidum* (DcPAL) were clustered in the same clade, whereas OsaPAL2
19 was alone in a neighboring branch suggesting it belongs to a different subfamily. In
20 addition, PALs of ferns and gymnosperms formed independent clusters, a result
21 consistent with the evolution of plants. PALs of plants and yeasts have a common
22 ancestor but are separated by a large evolutionary distance, with only 36% homology

1 between OsaPAL2 and RmPAL (CAA31486.1) from *Rhodotorula mucilaginosa*.

2

3 **3.4. Expression of recombinant proteins and catalytic product analysis**

4 The two ORFs were sub-cloned into *E.coli* vector pET-28a (+) by the In-fusion
5 method resulting in the heterologous plasmids pET28a-OsaPAL2 and
6 pET28a-OsaPAL62 (Fig. 2A). Recombinant His₆-tagged OsaPAL2 and OsaPAL62
7 were expressed in *E. coli* and purified on Ni-NTA resin. SDS-PAGE analysis showed
8 that both OsaPAL2 and OsaPAL62 were expressed and had molecular weights of
9 approximately 75 kDa (Fig. 2B).

10 Enzymes were incubated with L-Phe or L-Tyr and the products analyzed by
11 HPLC-UV, HPLC-MS, and NMR. The data showed that the products of L-Phe
12 catalyzed by OsaPAL2 and OsaPAL62 (OsaPAL2-L-Phe-p and OsaPAL62-L-Phe-p,
13 respectively), had the same UV spectra as that of *trans*-cinnamic acid. In HPLC-MS,
14 the most intense peaks of the two products were located at m/z 147 resulting from the
15 loss of a hydrogen radical. The structure of these two products was further confirmed
16 by NMR after semi-preparative isolation to be *trans*-cinnamic acid (Fig. 5).

17 For the corresponding products of L-Tyr, HPLC-UV analysis showed that both
18 enzymes converted L-Tyr to the new products, OsaPAL2-L-Tyr-p and
19 OsaPAL62-L-Tyr-p with the same UV-spectra as that of *p*-coumaric acid. Both
20 OsaPAL2-L-Tyr-p and OsaPAL62-L-Tyr-p exhibited an abundant ion at m/z 163
21 resulting from the loss of a hydrogen atom with subsequent loss of a carboxyl group
22 to form an ion at m/z 119. NMR analysis further confirmed OsaPAL2-L-Tyr-p and

1 OsaPAL62-L-Tyr-p to be *p*-coumaric acid (Fig. 6).

2 These results strongly suggest that OsaPAL2 and OsaPAL62 have both PAL and
3 TAL activity consistent with the activity of PALs present in other monocotyledonous
4 plants including *Zea mays*⁵¹ and *Phyllostachys edulis*.⁵²

6 3.5. Biochemical analysis of OsaPALs

7 In many previous works the formation of *trans*-cinnamic acid was usually
8 monitored at 275 or 280 nm.^{26,51,53} However, at this wavelength the substrate L-Phe
9 shows UV absorption (Fig. 5), which means that the determination of *trans*-cinnamic
10 acid will be disturbed when using a multimode reader. To eliminate the effect of the
11 substrate, we monitored absorbance at 300 nm.

12 The optimal pH and temperature for efficient enzyme activity of the purified
13 enzymes, OsaPAL2 and OsaPAL62, were found to be similar for L-Phe and L-Tyr
14 deamination (Fig. 7). Both enzymes functioned better in the pH range 8.5-10.0 and
15 showed only low activity at pH <7.0 or >11.0. This optimum pH range is comparable
16 to those of PALs isolated from other monocotyledonous species such as PePAL from
17 *Phyllostachys edulis* (pH 8.5–9.0),⁵² SmPAL1 from *Salvia miltiorrhiza* (pH 8.7)² and
18 BoPAL4 from *B. oldhamii* (pH 9.0).⁵⁰ The effect of temperature on the activity of
19 OsaPAL2 and OsaPAL62 was similar with both showing highest activity at 44°C.
20 Similar temperatures have been observed for maximum activity of PALs in *Rhus*
21 *chinensis* (45°C)¹ and *Arabidopsis* (46–48°C)⁴⁴ with a slightly lower temperature
22 being optimum for AvPAL in *A. variabilis* (40°C)²⁶ and a higher temperature being

1 optimum for Pal in *Helianthus annuus L* (55°C).⁵⁴

2 Assays were performed with L-Phe or L-Tyr as the substrate at optimum pH and
3 temperature to evaluate the kinetic constants of OsaPAL2 and OsaPAL62 (Figs. 7, 8
4 and Table 3). Both wild-type enzymes displayed marked kinetic preference for L-Phe
5 with a K_m of OsaPAL2 for L-Phe ($371.5 \pm 26.7 \mu\text{M}$), some 30-fold lower than that for
6 L-Tyr ($11690 \pm 1410 \mu\text{M}$), and a k_{cat}/K_m substantially higher for L-Phe ($2160 \text{ M}^{-1}\text{s}^{-1}$)
7 than for L-Tyr ($39.65 \text{ M}^{-1}\text{s}^{-1}$). A similar situation was found for OsaPAL62 with a K_m
8 for L-Phe ($593.1 \pm 48.6 \mu\text{M}$), about 60-fold lower than for L-Tyr ($39280 \pm 472 \mu\text{M}$),
9 and a k_{cat}/K_m value higher for L-Phe ($934.9 \text{ M}^{-1}\text{s}^{-1}$) than for L-Tyr ($16.69 \text{ M}^{-1}\text{s}^{-1}$).

10 Assays performed to determine the kinetic constants of OsaPAL2 and OsaPAL62 for
11 L-Tyr were unable to saturate the enzymes due to lower appetency and solubility
12 limitations for L-Tyr.⁴⁴ K_m values of these enzymes for L-Phe were similar to other
13 PALs, including the PAL of *Oryza sativa* ($500 \mu\text{M}$)⁵⁵ and the BoPAL2 of *B.oldhamii*
14 ($333 \mu\text{M}$).⁵⁰ k_{cat}/K_m values of these two wild-type enzymes for L-Phe were lower than
15 those of the PALs in *Zea mays* ($3.9 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$)⁵¹ and *Jatropha curcas L* ($1.384 \times$
16 $10^4 \text{ M}^{-1}\text{s}^{-1}$).⁵⁶

17 The higher L-Phe and lower L-Tyr deamination activity of OsaPAL2 and
18 OsaPAL62 is consistent with previous studies of PALs found in monocotyledons⁵¹.
19 Moreover, the OsaPALs were able to catalyze the reversible reaction whereby
20 *trans*-cinnamic acid is converted to L-Phe, an ability also found for some other PALs
21 (Data not shown).

22

1 3.6. Identification of specificity determining residues

2 OsaPAL2 and OsaPAL62 also have TAL activity allowing direct conversion of
3 L-Tyr to *p*-coumarate and thus eliminating the requirement for the hydroxylation step.
4 This activity is valuable for designing heterologous pathways for producing target
5 metabolites. The underlying mechanism of bi-functional PAL was considered as the
6 presence of a Phe residue controlling substrate selection.⁵⁷ Although several
7 experiments have supported this hypothesis, various questions remain unaddressed. In
8 a report by Watts *et al*,⁵⁷ when Phe144 residue was replaced by His in *A. thaliana*
9 PAL1, the mutant displayed a marked reduction (80-fold) in k_{cat}/K_m value, together
10 with an 18-fold increase in catalytic efficiency towards tyrosine. This result clearly
11 showed that Phe144 controlled substrate selection.⁵⁷ Hsieh *et al*,⁵⁰ however, did not
12 obtain the same results. When Phe133 was substituted with His, only a slightly
13 increased k_{cat}/K_m value toward tyrosine was observed in the F133H mutant. The
14 authors inferred that other residues also contribute to the substrate selectivity of PALs.
15 In view of this result, they proposed that further site-directed mutagenesis
16 experiments would be necessary for elucidation of the underlying mechanism of
17 substrate selectivity.⁵⁰ This situation poses the question of how, precisely, the Phe
18 residue functions in the substrate selectivity switch. Site-directed mutagenesis of Phe
19 residue was accordingly performed in OsaPALs with the aim of identifying this
20 function.

21 TAL and PAL have different substrate specificities associated with a difference of a
22 residue corresponding to F134 in OsaPAL2 and F128 in OsaPAL62 (Fig. 3). This

1 residue gives rise to low L-Tyr deamination activity which is improved by substituting
2 the two residues by H to give OsaPAL2F134H and OsaPAL62F128H (Table 3, Figs.
3 8-10). The optimum pH and temperature for the two mutants are almost the same as
4 the wild type proteins (Supplementary figs.1-3). The kinetic parameters, however,
5 varied markedly when F was mutated to H. For OsaPAL2F134H, values of K_m and k_{cat}
6 $/K_m$ for L-Tyr were 16.9-fold lower and 7.0-fold higher, respectively, than those for
7 the wild-type enzyme. After F→H mutation, the activity toward L-Phe decreased with
8 K_m being 3-fold higher than for OsaPAL2 and k_{cat}/K_m being reduced by 86% (Table 3).
9 OsaPAL62F128H exhibited marked TAL activity compared with the wild-type
10 enzyme (Table 3, Figs. 8,10) with values of K_m and k_{cat}/K_m for L-Tyr approximately
11 30-fold lower and 6-fold higher, respectively, than those of the wild-type enzyme.
12 Substitution of F with H resulted in greatly reduced activity for L-Phe with lower
13 affinity (K_m 2.8-fold higher than that of wild-type) and lower catalytic efficiency (k_{cat}
14 $/K_m$ 4.7-fold lower than that of the wild-type) (Table 3).

15 The fact that TAL activity of the OsaPALs is significantly increased in the F→H
16 mutations proves this F residue is associated with substrate specificity in agreement
17 with some previous studies.^{1,58} The compound, MIO, shown to be present in the
18 crystal structure of PAL,⁵⁹ is believed to be involved in the PAL reaction as a cofactor
19 attacking the carbon atom at the C2 position of the aromatic ring of L-Phe in a
20 Friedel–Crafts-type reaction.⁴⁷ H134 and H128 of OsaPAL2F134H and
21 OsaPAL62F128H, respectively, interact with the hydroxyl group of L-Tyr, resulting in
22 an increase in affinity.⁵⁸ The hydroxyl group of L-Tyr also acts as an

1 electron-donating group to increase the electron density at the C1,C3 and C5 positions
2 of the aromatic ring,⁶⁰ thereby interfering with the attack by MIO, so the mutants
3 retained higher catalytic activities towards L-Phe.⁶¹

4

5 **3.7. Residue substitution successfully improves enzymes activity**

6 The enzyme-kinetic properties of the purified proteins verified their biochemical
7 function as *bona fide* OsaPALs, a finding which is expected to clarify the biosynthetic
8 pathway of OSW-1. In addition, the successful characterization of the OsaPALs
9 provides useful information to reconstruct the OSW-1 biosynthetic pathway and allow
10 scale-up of the production of OSW-1 using more active enzymes.

11 After successful functional characterization, the two sequences were aligned with
12 three known PALs from yeasts, together with other three PALs from plants (Fig. 11).
13 The reason for selection of these PALs was that PALs from yeasts⁶¹⁻⁶³ appear to have
14 higher catalytic activity than those from plants.^{1,51,64} Examination of the sequence
15 alignment revealed 3 residues around ASG triad that were not absolutely conserved.
16 The three amino acids are T, V, and G in plant PALs. In yeast PALs, the equivalent
17 residues varied to S, S, and A. This is an interesting phenomenon which has not been
18 reported to date in the literature. To identify the function of the three amino acids,
19 site-directed mutagenesis was performed based on consensus approach.⁶⁵⁻⁶⁷

20 The mutants were OsaPAL2T196S and OsaPAL62T194S in which T196 of
21 OsaPAL2 and the T194 of OsaPAL62 were replaced with S; OsaPAL2V202S and
22 OsaPAL62V200S in which V202 of OsaPAL2 and V200 of OsaPAL62 were replaced

1 with S; and OsaPAL2G209A and OsaPAL62G207A in which G209 of OsaPAL2 and
2 G207 of OsaPAL62 were replaced with A. In addition, two double mutants,
3 OsaPAL2V202S/G209A and OsaPAL62V200S/G207A, and two triple mutants,
4 OsaPAL2T196S/V202S/G209A and OsaPAL62T194S/V200S/G207A, were
5 constructed.

6 After protein expression and purification (Fig. 12), the optimum pH and
7 temperature for these mutants were determined to be the same as the wild proteins
8 (Supplementary figs. 1-3). Then, the catalytic activities of these mutants were evaluated
9 (Table 3, Figs. 8-10) revealing that, although all the mutations lay near the PAL motif,
10 the effects were diverse (Table 3). Using L-Phe as a substrate, both T→S mutations
11 (OsaPAL2T196S and OsaPAL62T194S) and V→S mutations (OsaPAL2V202S and
12 OsaPAL62V200S) exhibited lower catalytic efficiency with lower values of k_{cat}/K_m
13 than those of the wild-type enzymes, in particular OsaPAL62V200S whose k_{cat}/K_m
14 value was only 0.33% of that of OsaPAL62.

15 Interestingly, the T→S mutation, where one hydrophilic amino acid replaces
16 another, showed lower affinity with K_m values 2-fold higher and lower catalytic
17 efficiency compared to the wild-type. In contrast, the V→S mutation, where a
18 hydrophilic amino acid replaces a hydrophobic one, showed higher affinity than the
19 wild-type. Furthermore, the G→A mutants were not significantly different in affinity
20 but exhibited higher catalytic efficiency than wild type.

21 The double mutants OsaPAL2V202S/G209A and OsaPAL62V200S/G207A
22 exhibited activities too low to be detected by the multimode reader but were shown by

1 HPLC analysis to have only 0.006% and 0.129% respectively of the wild-type
2 activities for L-Phe (Figs. 9, 10). The introduction of a third mutation to give the triple
3 mutants OsaPAL2T196S/V202S/G209A and OsaPAL62T194S/V200S/G207A
4 restored some degree of activity for L-Phe (0.016–0.339%) (Figs. 9, 10). The relative
5 activities of other mutants were also evaluated by HPLC and shown to be consistent
6 with their kinetic constants. In brief, the activity of OsaPAL2G209A and
7 OsaPAL62G207A was highest, followed by the wild-type, followed by
8 OsaPAL2T196S and OsaPAL62T194S with slightly lower activity than the wild-type
9 followed by OsaPAL2V202S and OsaPAL62V200S. Activities of the double mutants
10 OsaPAL2V202S/G209A and OsaPAL62V200S/G207A and triple mutants
11 OsaPAL2T196S/V202S/G209A and OsaPAL62T194S/V200S/G207A were negligible.
12 Similar results were obtained by HPLC for L-Tyr deamination activity (Figs. 9, 10).

13 In both of the two proposed reaction mechanisms for PAL,^{48,60,68,69} the A–S–G triad
14 in the motif can rearrange to a MIO ring to attack substrates. The importance of the
15 triad has been demonstrated in many studies but little research has focused on other
16 conserved amino acids around it.

17 The results described above indicate that residues adjacent to the A–S–G triad may
18 play different roles. In considering the K_m values of the single mutants in relation to
19 the positions of mutations, the higher affinity but lower catalytic efficiency of
20 OsaPAL2V202S and OsaPAL62V200S suggest the S202 in the former and S200 in
21 the latter may form hydrogen bonds with the β -proton of the substrate and hinder its
22 elimination. However, elucidation of the exact mechanism needs further investigation.

1 **4. Conclusions**

2 **4.1. First functional characterization of PAL cDNAs from Asparagaceae plant** 3 **species**

4 PALs are fairly ubiquitous in plants, including monocots, dicots, gymnosperms, ferns,
5 lycopods, liverworts, and algae. There is, however, no report about PALs from
6 Asparagaceae plant species.

7

8 **4.2. Reveal the function of three amino acids around ASG triad**

9 PALs from yeasts appear to have higher catalytic activity than those from plants.
10 Examination of the sequence alignment revealed 3 residues around ASG triad that was
11 not absolutely conserved. The three amino acids are T, V, and G in plant PALs. In
12 yeast PALs, the equivalent residues varied to S, S, and A. The amino acid difference
13 may be the reason of catalytic activity discrepancy between plant PALs and yeast
14 PALs. Mutagenesis studies demonstrated that the Thr-to-Ser mutants, OsaPAL2T196S
15 and OsaPAL62T194S, showed significantly higher substrate affinity than wild-types.
16 The Gly-to-Ala mutants, OsaPAL2G209A and OsaPAL62G207A, also showed higher
17 PAL and TAL activities. To the best of our knowledge, this is the first report about the
18 three mutants. It is also the first time that we have improved both PAL and TAL
19 activities by single amino acid substitutions.

20

21 **4.3. Further clarify the Phenylalanine function as a substrate selective switch**

22 There are some confusing results about Phe function as a substrate selective switch. In

1 the paper presented by Watts et al,⁵⁸ when Phe144 residue was replaced by His in
2 *A.thaliana* PAL1, the mutant displayed a marked reduction (80-fold) in k_{cat}/K_m value,
3 together with 18-fold increase in the catalytic efficiency towards tyrosine. This result
4 clearly showed Phe144 controlled the substrate selection. But this was not the case in
5 the report published by Hsieh et al.⁵¹ When Phe133 was substituted with His, slightly
6 increased k_{cat}/K_m value toward tyrosine was found in the F133H mutant. They
7 inferred other residues also contribute to the substrate selectivity of PALs. To further
8 clarify the phenylalanine function, two mutants OsaPAL2F134H and
9 OsaPAL62F128H were produced by site-directed mutagenesis. Results corroborated
10 Phe is indeed a control of substrate selection.

11

12 Acknowledgements

13 This work was supported by National Mega-project for Innovative Drugs (No.
14 2012ZX09301002), Open Foundation of State Key Laboratory of Bioactive Substance
15 and Function of Natural Medicines (B-2011-4) and PUMC Youth Fund (2012J21,
16 3332013112).

17

18 References

- 19 1 A. B. Oliveira, C. F. Moura, E. Gomes-Filho, C. A. Marco, L. Urban and M. R. Miranda, *PLoS One*,
20 **2013**, 8, e56354.
21 2 J. Song and Z. Wang, *Mol Biol Rep*, **2009**, 36, 939-952.
22 3 Y. Mimaki, M. Kuroda, A. Kameyama, Y. Sashida, T. Hirano, K. Oka, R. Maekawa, T. Wada, K.
23 Sugita and J. A. Beutler, *Bioorganic & Medicinal Chemistry Letters*, **1997**, 7, 633-636.
24 4 K. Tamura, H. Honda, Y. Mimaki, Y. Sashida and H. Kogo, *Br J Pharmacol*, **1997**, 121, 1796-1802.
25 5 Y. Zhou, C. Garcia-Prieto, D. A. Carney, R. H. Xu, H. Pelicano, Y. Kang, W. Yu, C. Lou, S. Kondo,
26 J. Liu, D. M. Harris, Z. Estrov, M. J. Keating, Z. Jin and P. Huang, *J Natl Cancer Inst*, **2005**, 97,

- 1 1781-1785.
- 2 6 S. Deng, B. Yu, Y. Lou and Y. Hui, *J Org Chem*, **1999**. 64, 202-208.
- 3 7 J. W. Morzycki and A. Wojtkielewicz, *Carbohydr Res*, **2002**. 337, 1269-1274.
- 4 8 M. Tsubuki, S. Matsuo and T. Honda, *Tetrahedron Letters*, **2008**. 49, 229-232.
- 5 9 W. Yu and Z. Jin, *J Am Chem Soc*, **2002**. 124, 6576-6583.
- 6 10 J. Xue, P. Liu, Y. Pan and Z. Guo, *J Org Chem*, **2008**. 73, 157-161.
- 7 11 S. Kubo, Y. Mimaki, M. Terao, Y. Sashida, T. Nikaido and T. Ohmoto, *Phytochemistry*, **1992**. 31,
8 3969-3973.
- 9 12 M. Kuroda, Y. Mimaki, A. Yokosuka, F. Hasegawa and Y. Sashida, *J Nat Prod*, **2002**. 65,
10 1417-1423.
- 11 13 M. Kuroda, Y. Mimaki, A. Yokosuka, Y. Sashida and J. A. Beutler, *J Nat Prod*, **2001**. 64, 88-91.
- 12 14 A. Podstolski, D. Havkin-Frenkel, J. Malinowski, J. W. Blount, G. Kourteva and R. A. Dixon,
13 *Phytochemistry*, **2002**. 61, 611-620.
- 14 15 D. Sircar and A. Mitra, *J Plant Physiol*, **2009**. 166, 1370-1380.
- 15 16 R. Loscher and L. Heide, *Plant Physiol*, **1994**. 106, 271-279.
- 16 17 J. Koukol and E. E. Conn, *J Biol Chem*, **1961**. 236, 2692-2698.
- 17 18 X. Hou, F. Shao, Y. Ma and S. Lu, *Mol Biol Rep*, **2013**. 40, 4301-4310.
- 18 19 C. J. Dong and Q. M. Shang, *Planta*, **2013**. 238, 35-49.
- 19 20 S. Giberti, C. M. Berteza, R. Narayana, M. E. Maffei and G. Forlani, *J Plant Physiol*, **2012**. 169,
20 249-254.
- 21 21 H. Yoon, Y. H. You, Y. E. Kim, Y. J. Kim, W. S. Kong and J. G. Kim, *J Microbiol Biotechnol*, **2013**.
22 23, 1055-1059.
- 23 22 C. A. Vaslet, R. L. Strausberg, A. Sykes, A. Levy and D. Filpula, *Nucleic acids research*, **1988**. 16,
24 11382.
- 25 23 C. W. Abell and R. S. Shen, *Methods Enzymol*, **1987**. 142, 242-253.
- 26 24 K. Kovacs, G. Banoczy, A. Varga, I. Szabo, A. Holczinger, G. Hornyanszky, I. Zagyva, C. Paizs, B.
27 G. Vertessy and L. Poppe, *PLoS One*, **2014**. 9, e85943.
- 28 25 J. S. Williams, M. Thomas and D. J. Clarke, *Microbiology*, **2005**. 151, 2543-2550.
- 29 26 M. C. Moffitt, G. V. Louie, M. E. Bowman, J. Pence, J. P. Noel and B. S. Moore, *Biochemistry*,
30 **2007**. 46, 1004-1012.
- 31 27 X. H. Wang, M. Gong, L. Tang, S. Zheng, J. D. Lou, L. Ou, J. Gomes-Laranjo and C. Zhang, *Mol*
32 *Biol Rep*, **2013**. 40, 97-107.
- 33 28 Q. Jin, Y. Yao, Y. Cai and Y. Lin, *PLoS One*, **2013**. 8, e62352.
- 34 29 Y. Jiang, B. Xia, L. Liang, X. Li, S. Xu, F. Peng and R. Wang, *Mol Biol Rep*, **2013**. 40, 2293-2300.
- 35 30 C. Fang, Y. Zhuang, T. Xu, Y. Li, Y. Li and W. Lin, *Journal of chemical ecology*, **2013**. 39,
36 204-212.
- 37 31 F. Xu, G. Deng, S. Cheng, W. Zhang, X. Huang, L. Li, H. Cheng, X. Rong and J. Li, *Molecules*,
38 **2012**. 17, 7810-7823.
- 39 32 U. R. Bagal, J. H. Leebens-Mack, W. W. Lorenz and J. F. Dean, *BMC Genomics*, **2012**. 13, 1-9.
- 40 33 W. H. Xiao, J. S. Cheng and Y. J. Yuan, *J Biotechnol*, **2009**. 139, 222-228.
- 41 34 R. W. Whetten and R. R. Sederoff, *Plant Physiol*, **1992**. 98, 380-386.
- 42 35 G. Hao, X. Du, F. Zhao and H. Ji, *Biotechnol Lett*, 32, 305-314.
- 43 36 M. R. Young, G. H. N. Towers and A. C. Neish, *Canadian Journal of Botany*, **1966**. 44, 341-349.
- 44 37 U. Czichi and H. Kindi, *Hoppe Seylers Z Physiol Chem*, **1975**. 356, 475-485.

- 1 38 I. A. Veliky and S. M. Martin, *Can J Microbiol*, **1970**. 16, 223-226.
- 2 39 J. Q. Kong, D. Lu and Z. B. Wang, *Molecules*, **2014**. 19, 1608-1621.
- 3 40 Y. Wei, J. Thompson and C. Floudas, *Proceedings of the Royal Society A: Mathematical, Physical*
4 *and Engineering Science*, **2012**. 468, 831-850.
- 5 41 M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F.
6 Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson and D. G. Higgins,
7 *Bioinformatics*, **2007**. 23, 2947-2948.
- 8 42 K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei and S. Kumar, *Mol Biol Evol*, **2011**. 28,
9 2731-2739.
- 10 43 A. Kumar and B. E. Ellis, *Plant physiology*, **2001**. 127, 230-239.
- 11 44 F. C. Cochrane, L. B. Davin and N. G. Lewis, *Phytochemistry*, **2004**. 65, 1557-1564.
- 12 45 R. Shi, C. M. Shuford, J. P. Wang, Y. H. Sun, Z. Yang, H. C. Chen, S. Tunlaya-Anukit, Q. Li, J.
13 Liu, D. C. Muddiman, R. R. Sederoff and V. L. Chiang, *Planta*, **2013**. 238, 487-497.
- 14 46 J. Kyte and R. F. Doolittle, *Journal of molecular biology*, **1982**. 157, 105-132.
- 15 47 L. Poppe, *Curr Opin Chem Biol*, **2001**. 5, 512-524.
- 16 48 L. Poppe and J. Reteý, *Angew Chem Int Ed Engl*, **2005**. 44, 3668-3688.
- 17 49 J. Reteý, *Biochim Biophys Acta*, **2003**. 1647, 179-184.
- 18 50 L. S. Hsieh, G. J. Ma, C. C. Yang and P. D. Lee, *Phytochemistry*, **2010**. 71, 1999-2009.
- 19 51 J. Rosler, F. Krekel, N. Amrhein and J. Schmid, *Plant physiology*, **1997**. 113, 175-179.
- 20 52 Z. M. Gao, X. C. Wang, Z. H. Peng, B. Zheng and Q. Liu, *Plant Cell Rep*, **2012**. 31, 1345-1356.
- 21 53 J. D. Cui, L. L. Li and H. J. Bian, *PloS one*, **2013**. 8, e80581.
- 22 54 J. Jorriñ, R. Lopez-Valbuena and M. Tena, *Biochemistry international*, **1991**. 24, 1-11.
- 23 55 A. D. Sarma and R. Sharma, *Phytochemistry*, **1999**. 50, 729-737.
- 24 56 J. Gao, S. Zhang, F. Cai, X. Zheng, N. Lin, X. Qin, Y. Ou, X. Gu, X. Zhu, Y. Xu and F. Chen, *Mol*
25 *Biol Rep*, **2012**. 39, 3443-3452.
- 26 57 K. T. Watts, B. N. Mijts, P. C. Lee, A. J. Manning and C. Schmidt-Dannert, *Chem Biol*, **2006**. 13,
27 1317-1326.
- 28 58 G. V. LOUIE, M. E. BOWMAN, M. C. MOFFITT, T. J. BAIGA, B. S. MOORE and J. P. NOEL,
29 *Chem Biol*, **2006**. 13, 1327-1338.
- 30 59 J. C. Calabrese, D. B. Jordan, A. Boodhoo, S. Sariaslani and T. Vannelli, *Biochemistry*, **2004**. 43,
31 11403-11416.
- 32 60 B. Schuster and J. Réteý, *Proceedings of the National Academy of Sciences*, **1995**. 92, 8433-8437.
- 33 61 L. Zhu, W. Cui, Y. Fang, Y. Liu, X. Gao and Z. Zhou, *Biotechnol Lett*, **2013**. 35, 751-756.
- 34 62 I. L. Bazukian, A. E. Vardanian, A. A. Ambartsumian, P. V. Tozalakian and G. Popov Iu, *Prikl*
35 *Biokhim Mikrobiol*, **2009**. 45, 23-27.
- 36 63 C. T. Evans, K. Hanna, D. Conrad, W. Peterson and M. Misawa, *Applied microbiology and*
37 *biotechnology*, **1987**. 25, 406-414.
- 38 64 L. S. Hsieh, C. S. Yeh, H. C. Pan, C. Y. Cheng, C. C. Yang and P. D. Lee, *Protein Expr Purif*, **2010**.
39 71, 224-230.
- 40 65 M. Lehmann and M. Wyss, *Curr Opin Biotechnol*, **2001**. 12, 371-375.
- 41 66 R. A. Chica, N. Doucet and J. N. Pelletier, *Curr Opin Biotechnol*, **2005**. 16, 378-384.
- 42 67 S. Lutz, *Curr Opin Biotechnol*, **2010**. 21, 734-743.
- 43 68 K. R. Hanson and E. A. Havir, *Arch Biochem Biophys*, **1970**. 141, 1-17.
- 44 69 J. D. Hermes, P. M. Weiss and W. W. Cleland, *Biochemistry*, **1985**. 24, 2959-2967.

1 **Figure legends:**

2 Figure 1: (A) The structure of OSW-1 with the *p*-methoxybenzoyl group (dashed
3 circle) and the enzymes responsible for OSW-1 biosynthesis. (B) The proposed
4 pathways for the biosynthesis of the *p*-methoxybenzoyl group. Dashed circles show
5 the β -oxidation pathway and dashed boxes show the non β -oxidation pathway. Solid
6 arrows show the established biochemical steps and broken arrows the hypothetical
7 steps. Stacked arrows indicate the involvement of multiple enzymatic reactions. (PAL,
8 phenylalanine ammonia-lyase; C4H, cinnamate-4-hydroxylase; TAL, tyrosine
9 ammonia-lyase; CA, *trans*-cinnamic acid; pCA, *p*-coumaric acid).

10

11 Figure 2: cDNA cloning and wild-type enzyme expression.

12 (A) cDNA cloning of *OsaPAL2* and *OsaPAL62* and the construction of expression
13 plasmids.

14 (B) SDS-PAGE analysis of recombinant wild-type enzymes. From right to left; empty
15 vector as the control, *OsaPAL2* crude lysate, purified wild-type *OsaPAL2*, protein
16 standards, purified wild-type *OsaPAL62*, *OsaPAL62* crude lysate, empty vector and
17 molecular weights corresponding to protein standards. The arrows indicate the target
18 proteins.

19

20 Figure 3: Amino acid sequence alignment of *OsaPAL2* and *OsaPAL62*.

21 Asterisks, dots and colons indicate identical, semi-conserved and conserved
22 substitutions, respectively. The open and filled *triangles* indicate the residues

1 individually substituted in this study and the filled *triangles* also indicate the positions
2 of residues believed to be responsible for substrate selectivity. *Rectangles* and *arrows*
3 present the predicted α -helices and β -strands, respectively. The highly conserved
4 phenylalanine ammonia-lyase signature is circled and the conserved active site motif
5 (A-S-G) is indicated by the bracket.

6

7 Figure 4: Phylogenetic analysis of OsaPAL2, OsaPAL62 and other PALs.

8 The phylogenetic tree was constructed using the neighbor-joining method in the
9 MEGA5.1 program. It shows a comparison of the sequences from dicotyledons:
10 *Paeonia lactiflora* PIPAL (AFI71896.1), *Rhus chinensis* RcPAL (AGH13333.1),
11 *Populus trichocarpa* PtPAL (ACC63887.1), *Hibiscus cannabinus* HcPAL
12 (AFN85669.1), *Trifolium pratense* TrPAL (AAZ29733.1), *Scutellaria baicalensis*
13 SbPAL (ADN32768.1), *Melissa officinalis* MoPAL (CBJ23826.1), *Salvia miltiorrhiza*
14 SmPAL (ABD73282.1), *Quercus suber* QsPAL (AAR31107.1), *Rubus idaeus*
15 RiPAL (AAF40224.1), *Platycodon grandiflorus* PgPAL (AEM63671.1) and *Coffea*
16 *canephora* CcPAL (AEO92028.1); the monocotyledons: *Musa balbisiana* MbPAL
17 (BAG70992.1), *Lycoris radiata* LrPAL (ACM61988.1) and *Dendrobium candidum*
18 DcPAL (AGC23439.1); the gymnosperms: *Ginkgo biloba* GbPAL (ABU49842.1) and
19 *Ephedra sinica* EsPAL (BAG74772.1); the ferns: *Blechnum spicant* BsPAL
20 (AAW80643.1), *Equisetum arvense* EaPAL (AAW80639.1) and *Ophioglossum*
21 *reticulatum* OrPAL (AAW80642.1); the yeasts: *Rhodotorula mucilaginosa* RmPAL
22 (CAA31486.1), *Rhodosporidium toruloides* RtPAL2 (AAA33883.1), *Rhodotorula*

1 *glutinis* ATCC 204091 RgPAL2 (EGU13302.1), *Rhodotorula glutinis* RgPAL1
2 (ABB04148.1) and *Rhodospiridium toruloides* RtPAL1 (CAA35886.1).

3

4 Figure 5: Analysis of the products of OsaPALs *in vitro* assays with L-Phe.

5 (A) HPLC analysis: Traces 1 and 4 show the substrate (L-Phe) and product standard
6 (*trans*-cinnamic acid), respectively; traces 2 and 3 are the controls showing the results
7 of boiled enzymes incubated with substrate (L-Phe); Traces 5 and 6 are the results of
8 purified wild-type enzymes incubated with L-Phe and the resulted products have
9 similar UV-spectra with *trans*-cinnamic acid.

10 (B) Mass spectra (MS) analysis; both OsaPAL2-L-Phe-p and OsaPAL62-L-Phe-p
11 exhibit strongest ion peaks at 147 *m/z*.

12 (C) NMR analysis results indicate OsaPAL2-L-Phe-p and OsaPAL62-L-Phe-p have
13 the same structure.

14

15 Figure 6: Analysis of the products of OsaPALs *in vitro* assays with L-Tyr.

16 (A) HPLC analysis: Traces 1 and 4 show the substrate (L-Tyr) and product standard
17 (*p*-coumaric acid), respectively; traces 2 and 3 are the controls showing the results of
18 boiled enzymes incubated with substrate (L-Tyr); Traces 5 and 6 are the results of
19 purified wild-type enzymes incubated with L-Tyr and the resulted products have
20 similar UV-spectra with *p*-coumaric acid (Note the wavelength was set at 275 nm to
21 visualize the substrate L-Tyr).

22 (B) Mass spectra (MS) analysis; both OsaPAL2-L-Tyr-p and OsaPAL62-L-Tyr-p have

1 the same peaks at m/z 163 and 119.

2 (C) NMR analysis results and the structure of the product.

3

4 Figure 7: Biochemical characterization and comparison of recombinant enzymes.

5 The optimum temperatures and pH for catalyzing L-Phe (A, C) and L-Tyr (B, D) of
6 both enzymes were determined. (E) and (F) are the overlays of substrate saturation
7 plots of the two enzymes using L-Phe and L-Tyr as substrates performed in CHES
8 buffer (0.1 M, pH 9.5) at 44°C by continuous monitoring of the changes in
9 absorbance at 300 and 310 nm, respectively. (Note at a wavelength of 275 nm, the
10 substrate L-Phe shows UV absorption which was eliminated by monitoring
11 absorbance at 300 nm).

12

13 Figure 8: Overlays of substrate saturation plots of different enzymes.

14 (A) (B) Plots of initial velocity against L-Phe concentration for reactions of wild-type
15 and mutant enzymes.

16 (C) (D) Plots of initial velocity against L-Tyr concentration for reactions of wild-type
17 and mutant enzymes. Wild-type enzymes did not reach enzyme saturation due to the
18 limited solubility of L-Tyr.

19

20 Figure 9: Comparison of OsaPAL2 mutant and wild-type activities.

21 (A) (D) HPLC was used to compare OsaPAL2 wild-type and mutant activities using
22 the same amount of purified enzyme and substrate L-Phe or L-Tyr in an aqueous

1 buffer. This shows one of the results of a reaction lasting 3 h, the various heights of
2 products indicating the different amounts. (Note UV absorption of the product of
3 L-Phe was high at 275 nm producing flat-topped peaks; to improve the precision of
4 the results, we also monitored at 300 nm).

5 (B) (E) Partially enlarged views of 2V202S/G209A and 2T196S/V202S/G209A.

6 (C) (F) Relative activity was determined based on the average rates of 9 independent
7 experiments; scale was set to 100% for wild-type OsaPAL2. Results are presented as
8 means \pm SD.

9

10 Figure 10: Comparison of OsaPAL62 mutant and wild-type activities.

11 (A) (D) HPLC was used to compare OsaPAL62 wild-type and mutant activities
12 using the same amount of purified enzymes and substrate L-Phe or L-Tyr in an
13 aqueous buffer. This shows one of the results of a reaction lasting 3 h, the various
14 heights of products indicating the different amounts. (Note UV absorption of the
15 product of L-Phe was high at 275nm producing flat-topped peaks; to improve the
16 precision of the results, we also monitored at 300 nm).

17 (B) (E) Partially enlarged views of 62V200S, 62V200S/G207A and
18 62T194S/V200S/G207A.

19 (C) (F) Relative activity is presented as means \pm SD of 9 independent experiments,
20 scale was set to 100% for wild-type OsaPAL62.

21

22 Figure 11: Multiple alignment of PAL signatures of plants and yeasts.

1 Red circles highlight the active site motif (A-S-G) which undergoes autocatalysis to
2 form MIO. The residues that differ in yeasts and plants are in black circles. The
3 sequences used are OsaPAL2, OsaPAL62, *Rhus chinensis* RcPAL (AGH13333), *Zea*
4 *mays* ZmPAL (NP_001105334), *Bambusa oldhamii* BoPAL2 (ACN62413.1),
5 *Rhodotorula glutinis JN-1* RgPAL (see reference⁶¹), *Rhodospiridium toruloides*
6 RtPAL3 (P11544), *Rhodospiridium toruloides* RtPAL2 (AAA33883.1).

7

8 Figure 12: SDS-PAGE analysis of purified mutants of OsaPAL2 and OsaPAL62.

9 M stands for protein standards with corresponding molecular weights shown at right;
10 small black arrows indicate target proteins with similar molecular weights.

11

12

13

14

15

16

17

18

19

20

21

22

1 **Tables:**

2 Table 1: Primers used in gene cloning and plasmid construction

Primers	Sequences (5'-3')	Description
FPAL1	5'-CATCATAATCTGACGGTTTTTC-3'	Forward primer used for <i>OsaPAL2</i> amplification in the first round
FPAL2	5'-ATGGAGAACGGCAACGGTAA-3'	Forward primer used for <i>OsaPAL2</i> amplification in the second round
RPAL3	5'-CAGAATTATGAAATTCAGCC-3'	Reverse primer used for <i>OsaPAL2</i> amplification in the first round
RPAL4	5'-TCAACATATTGGCAGCGGTGC-3'	Reverse primer used for <i>OsaPAL2</i> amplification in the second round
FET28aPAL2	5'-TCGCGGATCCGAATTCATGGAGAACGGCA ACGGTAAC-3'	Forward primer used for pET28a- <i>OsaPAL2</i> construction
RET28aPAL2	5'-GTGCGGCCGCAAGCTTCAACATATTGGCA GCGGTGC-3'	Reverse primer used for pET28a- <i>OsaPAL2</i> construction
FPAL5	5'-CAATCAGCCGTTTACGAGACC-3'	Forward primer used for <i>OsaPAL62</i> amplification in the first round
FPAL6	5'-ATGGAATCCCTCCACGCCAAC-3'	Forward primer used for <i>OsaPAL62</i> amplification in the second round
RPAL1	5'-CCGAAGTACTGAATGAAAATC-3'	Reverse primer used for <i>OsaPAL62</i> amplification in the first round
RPAL2	5'-CTAGCAAATGGCAGGGGAG-3'	Reverse primer used for <i>OsaPAL62</i> amplification in the second round
FET28aPAL62	5'-TCGCGGATCCGAATTCATGGAATCCCTCCA CGCCAAC-3'	Forward primer used for pET28a- <i>OsaPAL62</i> construction
RET28aPAL62	5'-GTGCGGCCGCAAGCTTCTAGCAAATGGGCA GGGGAG-3'	Reverse primer used for pET28a- <i>OsaPAL62</i> construction

3

4

5

6

7

8

9

10

11

12

1 Table 2: Oligonucleotide primer pairs used for site-directed mutagenesis in this study.

2 The bases underlined in each primer are the codons introduced.

Mutant	Template	Primer	Sequence
OsaPAL2F134H	pET28a-OsaPAL2	PAL2F134HF	5'-TGCAGAGAGAGTTGATAAAGGC <u>AT</u> CTGAATGCCGG-3'
		PAL2F134HR	5'- <u>ATGC</u> CTTATCAACTCTCTCTGCAGGGCGCCCC-3'
OsaPAL2T196S	pET28a-OsaPAL2	PAL2T196SF	5'-TACGAGGCACCATC <u>AGCG</u> CCTCCGG-3'
		PAL2T196SR	5'- <u>GCT</u> GATGGTGCCTCGTAGAGGAAGGCA-3'
OsaPAL2V202S	pET28a-OsaPAL2	PAL2V202SF	5'-GCCTCCGGCGACCTA <u>AGC</u> CCCTTGTCT-3'
		PAL2V202SR	5'- <u>GCT</u> TAGGTCGCCGAGGCGGTGATGGTG-3'
OsaPAL2G209A	pET28a-OsaPAL2	PAL2G209AF	5'-TTGTCCTACATTGCC <u>GCG</u> CTTCTCACCG-3'
		PAL2G209AR	5'- <u>CGCG</u> GCAATGTAGGACAAGGGGACTAGG-3'
OsaPAL2V202S/G209A	pET28a-OsaPAL2	PAL2G209AF	5'-TTGTCCTACATTGCC <u>GCG</u> CTTCTCACCG-3'
		PAL2DMR	5'- <u>CGCG</u> GCAATGTAGGACAAGGG <u>GCT</u> TAGG-3'
OsaPAL2T196S/V202S/G209A	pET28a-OsaPAL2V202S/G209A	PAL2T196SF	5'-TACGAGGCACCATC <u>AGCG</u> CCTCCGG-3'
		PAL2T196SR	5'- <u>GCT</u> GATGGTGCCTCGTAGAGGAAGGCA-3'
OsaPAL62F128H	pET28a-OsaPAL62	PAL62F128HF	5'-TTCAGAAGGAGCT <u>CAT</u> CAGACATCTCAACGCGGG-3'
		PAL62F128HR	5'- <u>ATGT</u> CTGATGAGCTCCTTCTGAAGGGCACCACC-3'
OsaPAL62T194S	pET28a-OsaPAL62	PAL62T194SF	5'-CTCCGGGCACGATC <u>AGCG</u> CCTCCGGCG-3'
		PAL62T194SR	5'- <u>GCT</u> GATCGTGCCGGAGAGGGAGGCAC-3'
OsaPAL62V200S	pET28a-OsaPAL62	PAL62V200SF	5'-GCCTCCGGCGACCTA <u>AGC</u> CCGTTGTCT-3'
		PAL62V200SR	5'- <u>GCT</u> GAGGTCGCCGAGGCGGTGATCGTG-3'
OsaPAL62G207A	pET28a-OsaPAL62	PAL62G207AF	5'-ATATCGCC <u>GCG</u> ATCCTCACCG-3'
		PAL62G207A R	5'- <u>CGCG</u> GCGATATAGGACAACGG-3'
OsaPAL62V200S/G207A	pET28a-OsaPAL62	PAL62DMF	5'-TTGTCCTATATCGCC <u>GCG</u> ATCCTCACCG-3'
		PAL62DMR	5'- <u>CGCG</u> GCGATATAGGACAACGG <u>GCT</u> GAGG-3'
OsaPAL62T194S/V200S/G207A	pET28a-OsaPAL62V200S/G207A	PAL62T194SF	5'-CTCCGGGCACGATC <u>AGCG</u> CCTCCGGCG-3'
		PAL62T194SR	5'- <u>GCT</u> GATCGTGCCGGAGAGGGAGGCAC-3'

3

4

5

6

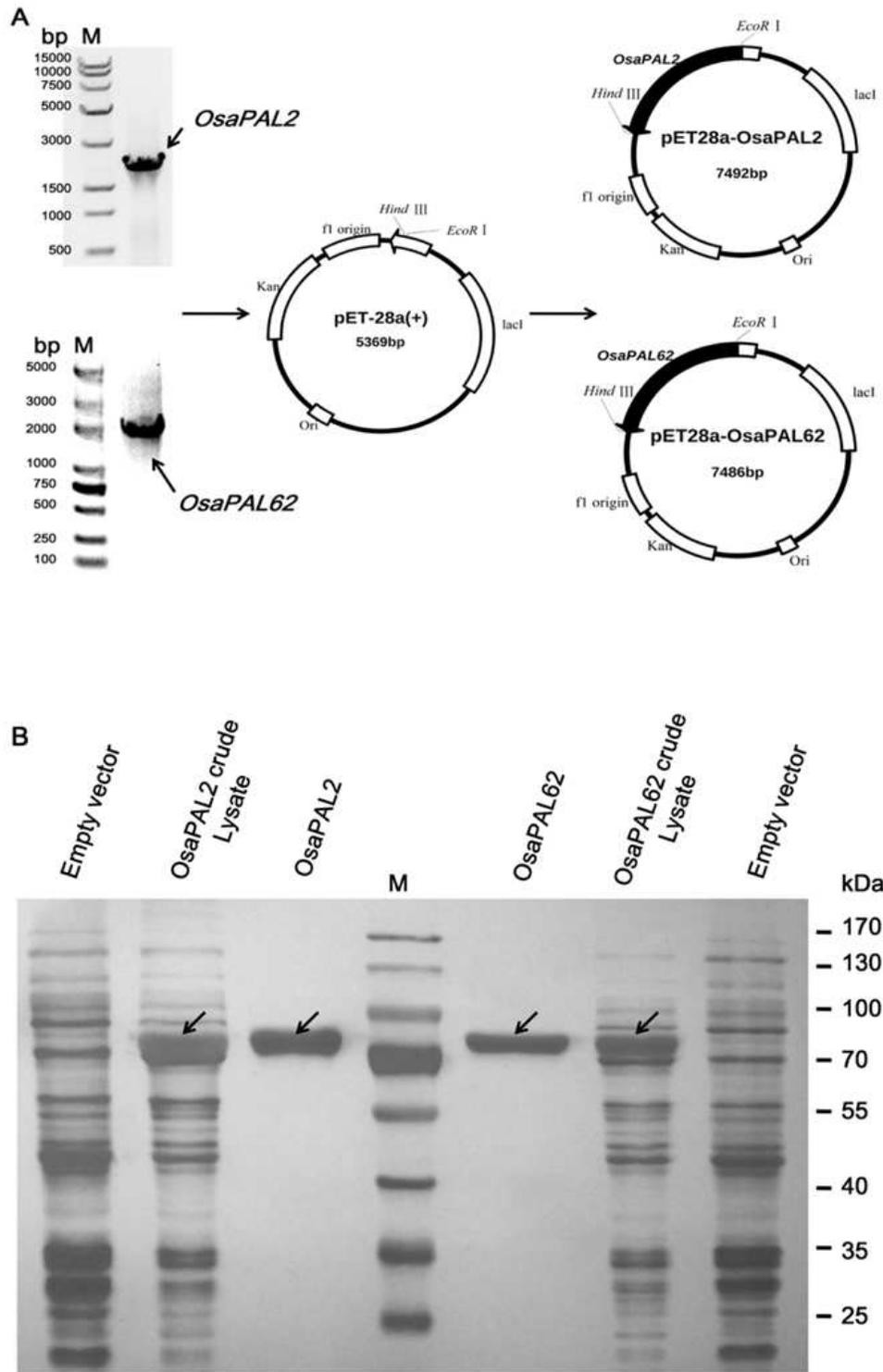
7

8

1 Table 3: Kinetic constants of OsaPAL2, OsaPAL62 and their mutants with L-Phe and
 2 L-Tyr as substrates. All assays were performed in aqueous 100 mM CHES (pH 9.5) at
 3 44°C and the formation of *trans*-cinnamic acid and *p*-coumric acid were monitored
 4 continuously by multimode reader at 300 and 310 nm, respectively.

Substrate	Enzyme	V _{max} (n k_{cat} / μ g protein)	K_m (μ M)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
L-Phe	OsaPAL2	51.48 \pm 0.82	371.5 \pm 26.7	0.8025	2160
	OsaPAL2F134H	23.40 \pm 0.23	1241 \pm 44	0.3646	293.8
	OsaPAL2T196S	33.78 \pm 0.36	662.9 \pm 28.8	0.5264	794.1
	OsaPAL2V202S	12.21 \pm 0.17	311.7 \pm 21.1	0.1903	610.5
	OsaPAL2G209A	119.3 \pm 2.2	790.4 \pm 57.54	1.861	2354
	OsaPAL62	35.82 \pm 0.70	593.1 \pm 48.6	0.5545	934.9
	OsaPAL62F128H	21.5 \pm 0.2	1679 \pm 43	0.3328	198.2
	OsaPAL62T194S	52.27 \pm 0.46	1236 \pm 39	0.8105	655.7
	OsaPAL62V200S	0.08034 \pm 0.00130	118.9 \pm 10.4	0.0003105	2.612
	OsaPAL62G207A	60.26 \pm 0.61	529.2 \pm 23.0	0.9332	1763
L-Tyr	OsaPAL2	29.73 \pm 2.41	11690 \pm 1410	0.4635	39.65
	OsaPAL2F134H	12.35 \pm 0.23	691.0 \pm 46.0	0.1925	278.6
	OsaPAL2G209A	30.08 \pm 1.33	3538 \pm 365	0.4690	132.6
	OsaPAL62	42.35 \pm 4.40	39280 \pm 472	0.6557	16.69
	OsaPAL62F128H	8.603 \pm 0.161	1320 \pm 74	0.1332	100.9
	OsaPAL62G207A	20.29 \pm 1.56	9423 \pm 1148	0.3143	33.35

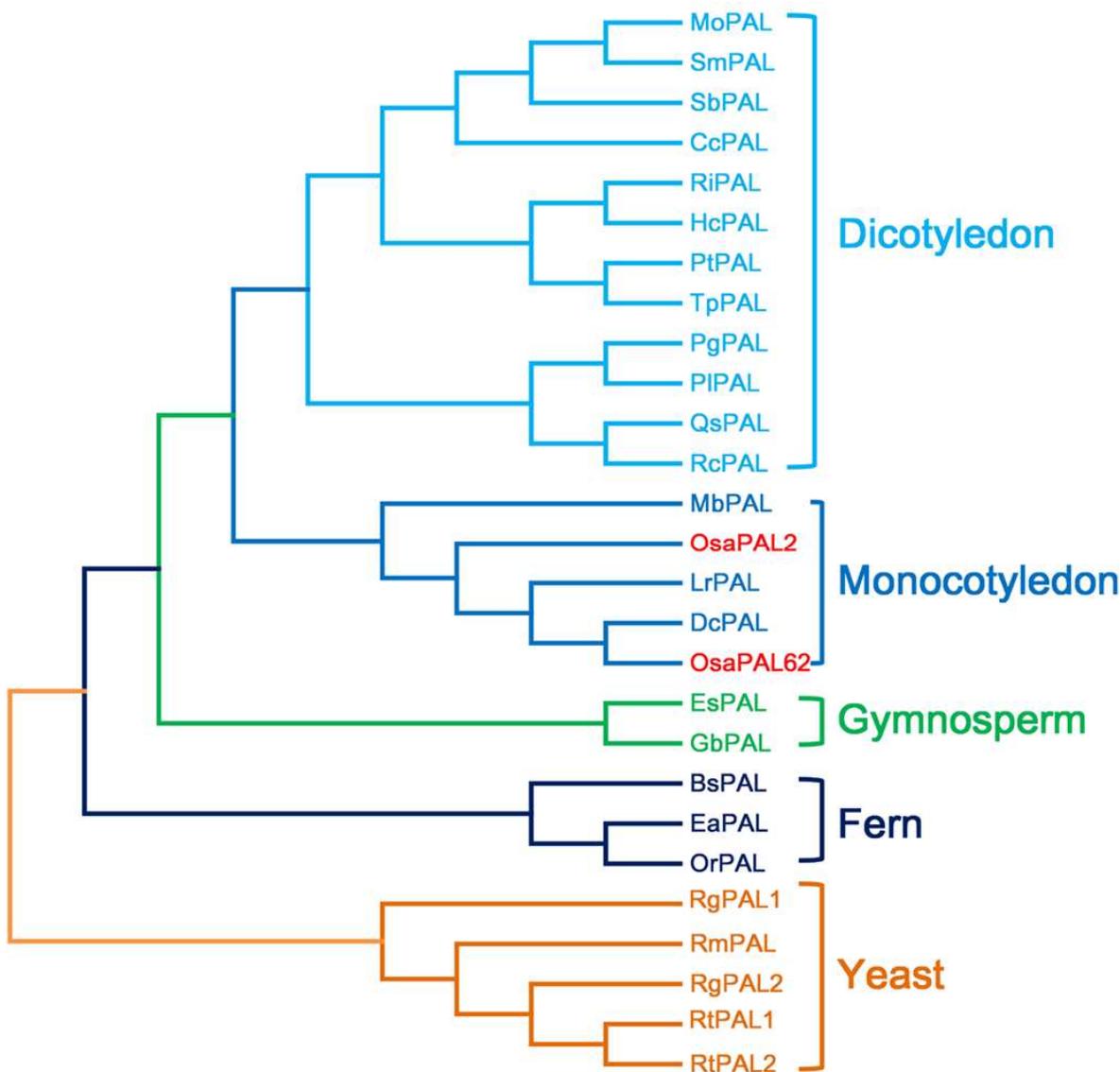
5



1

2

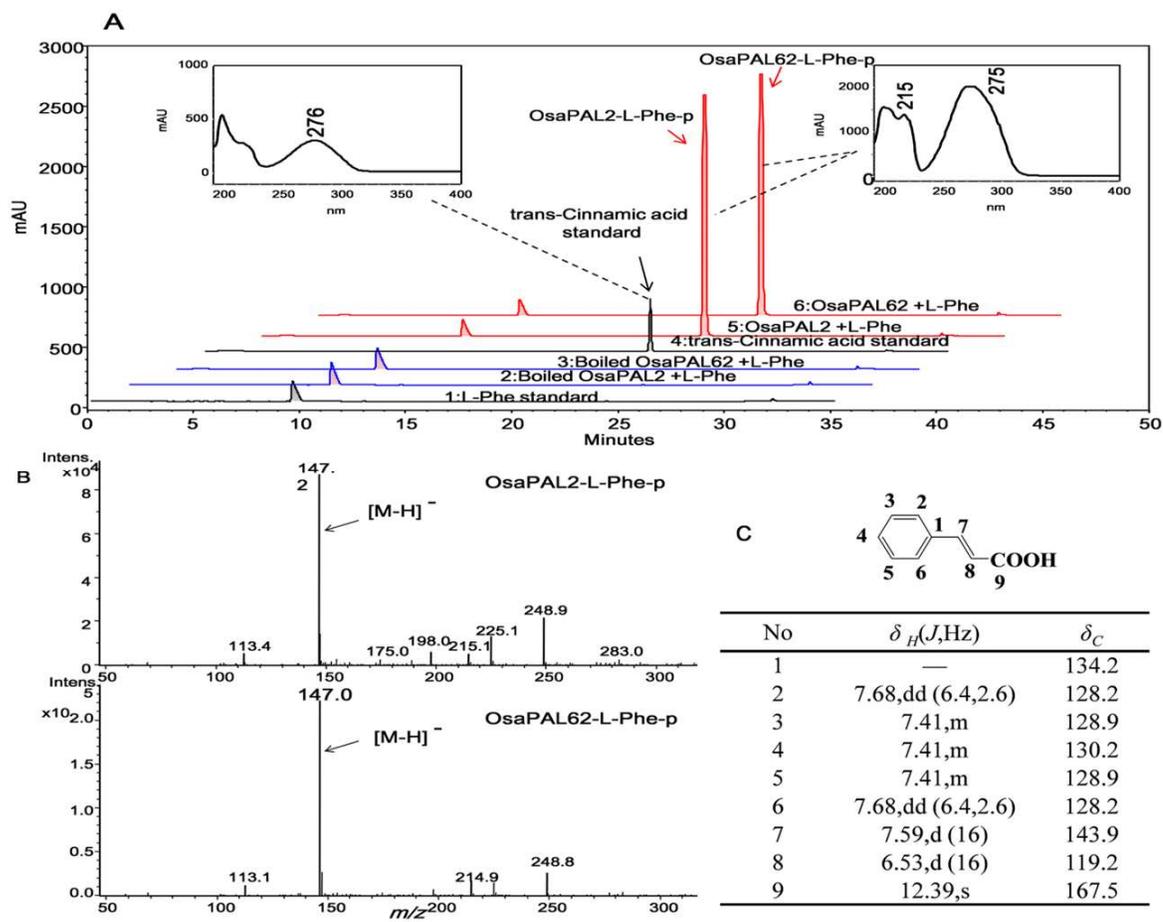
Figure 2



1
2

Figure 4

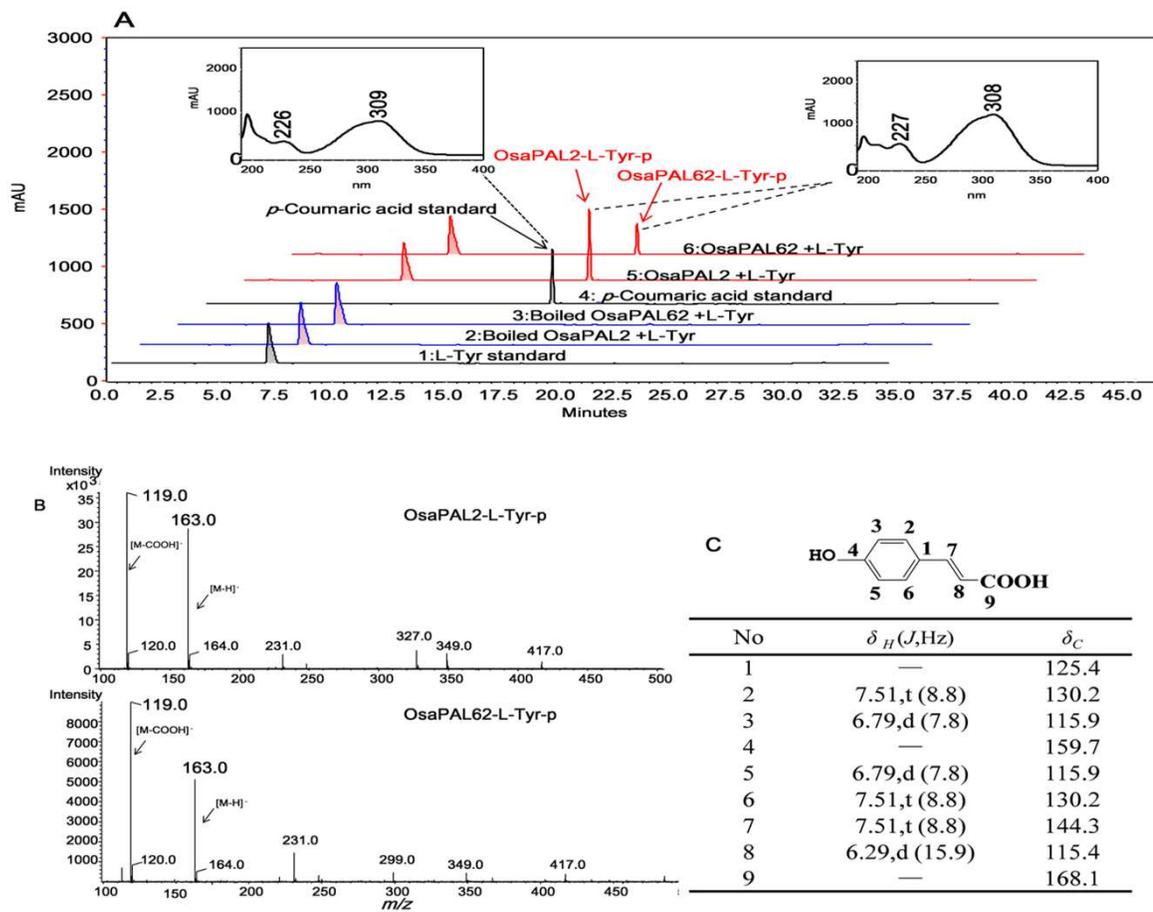
RSC Advances Accepted Manuscript



1

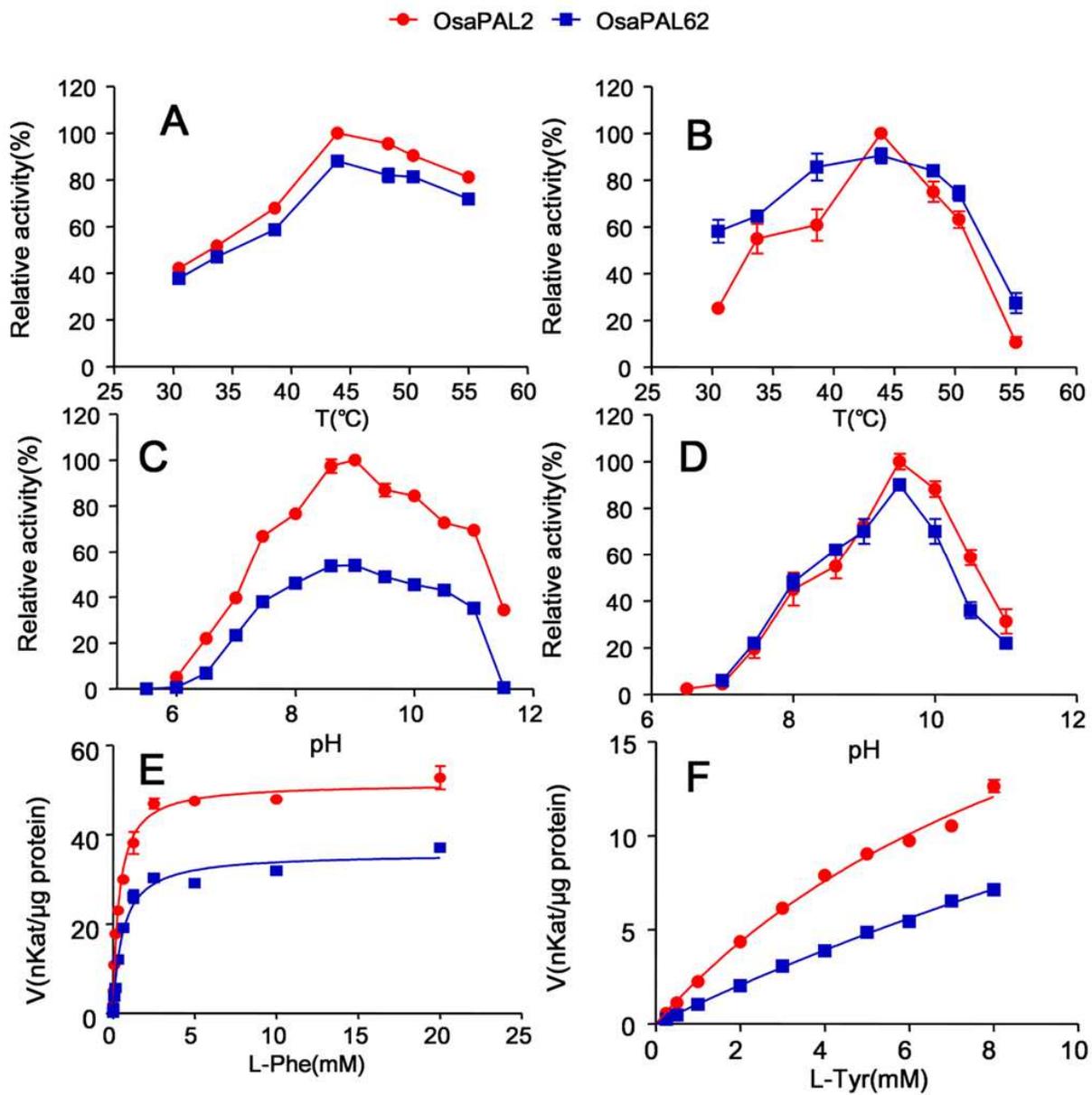
2

Figure 5



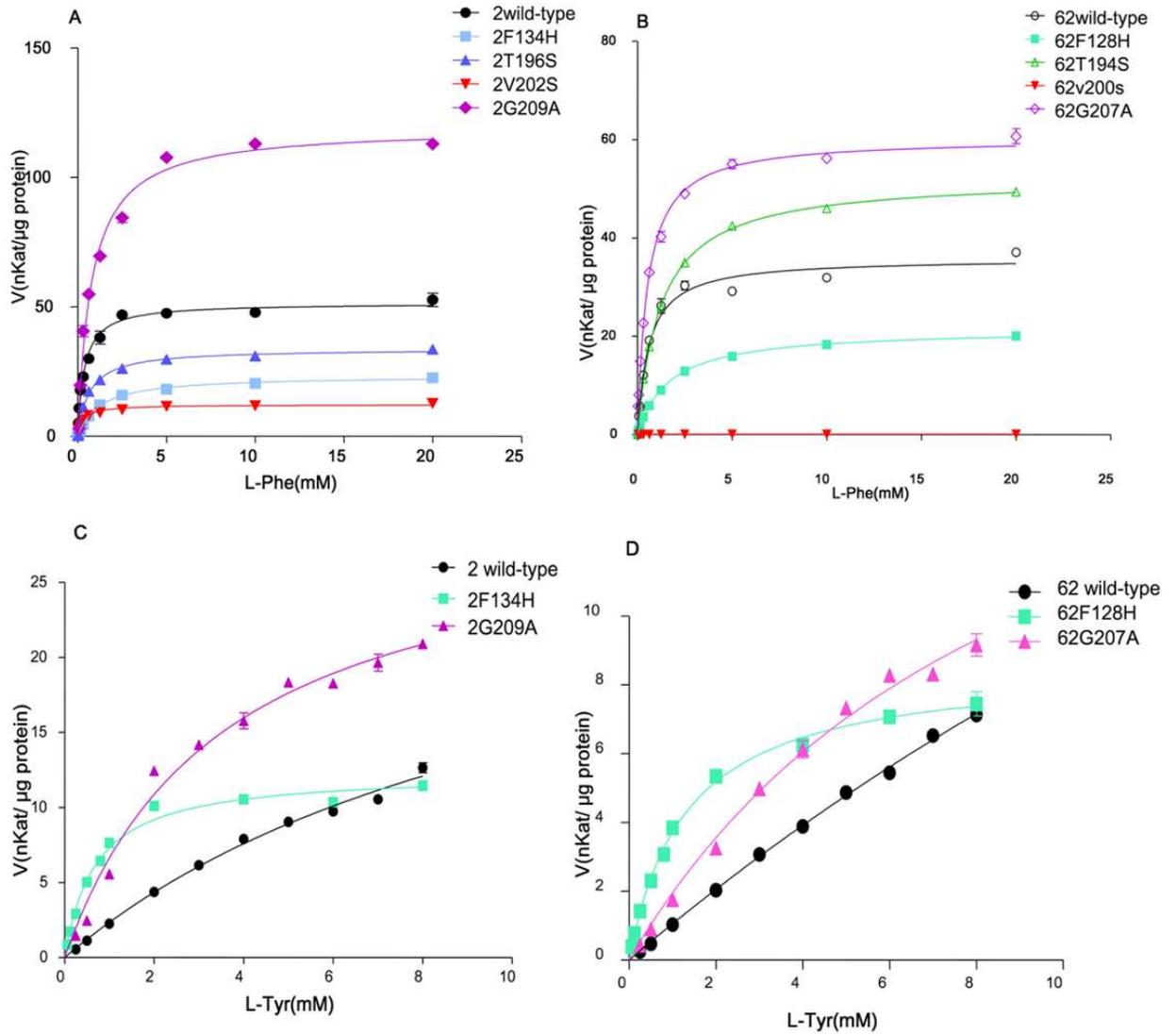
1
2

Figure 6



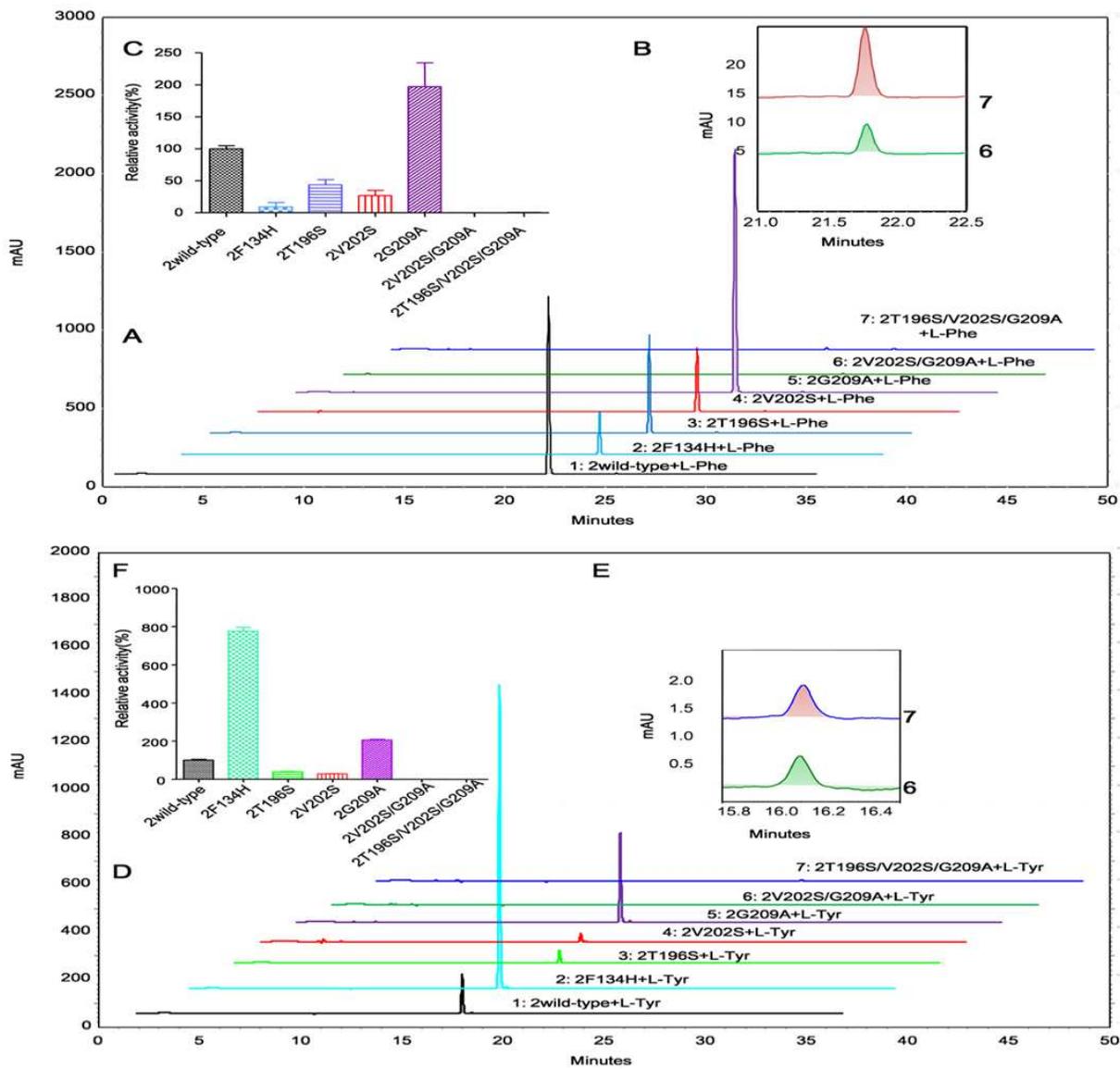
1
2

Figure 7



1
2
3

Figure 8



1
2

Figure 9

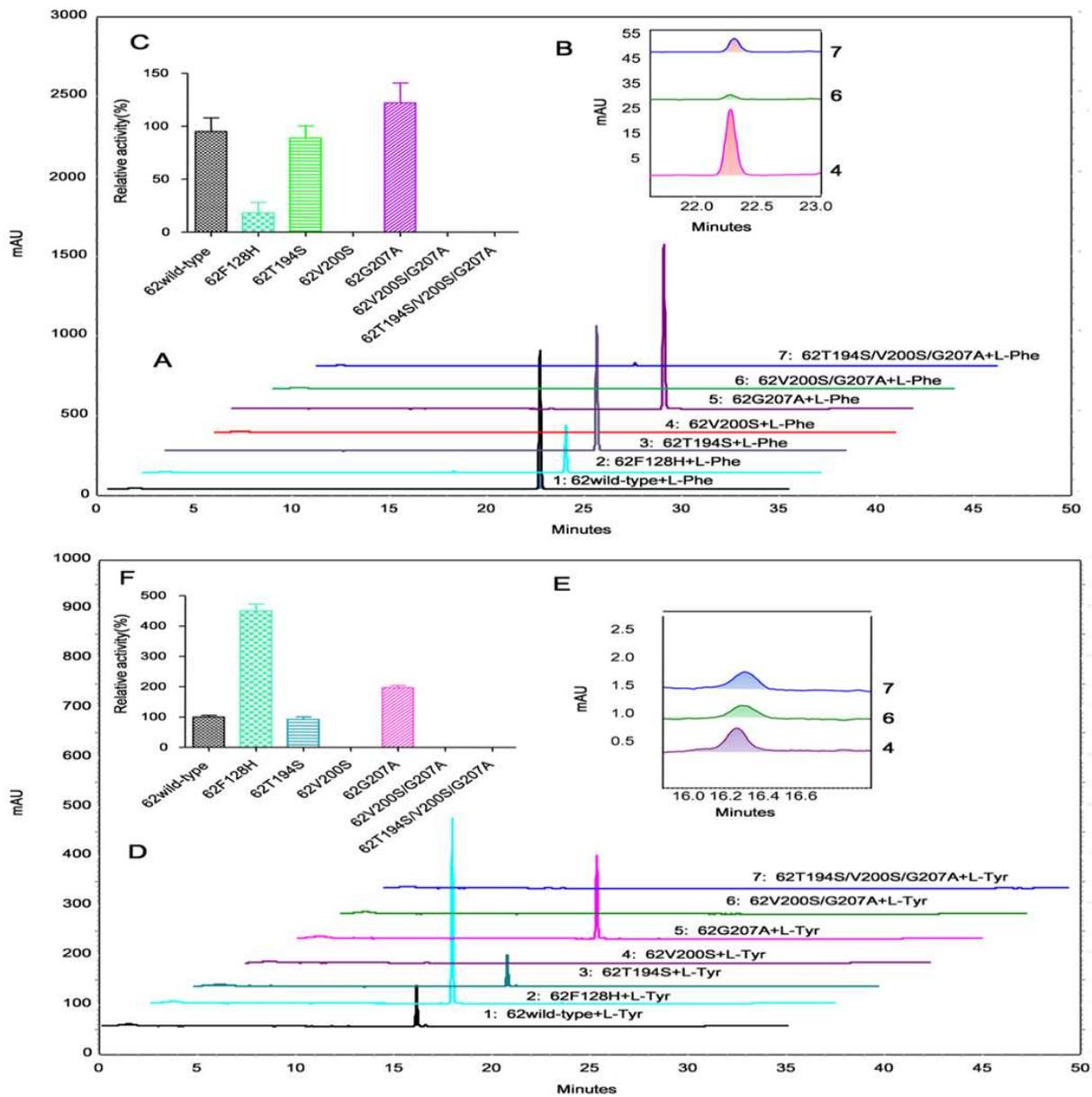
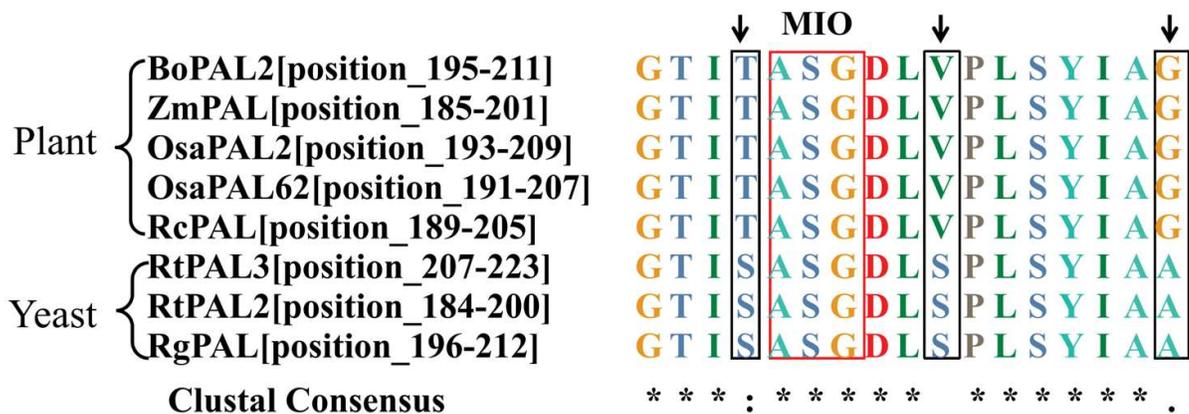


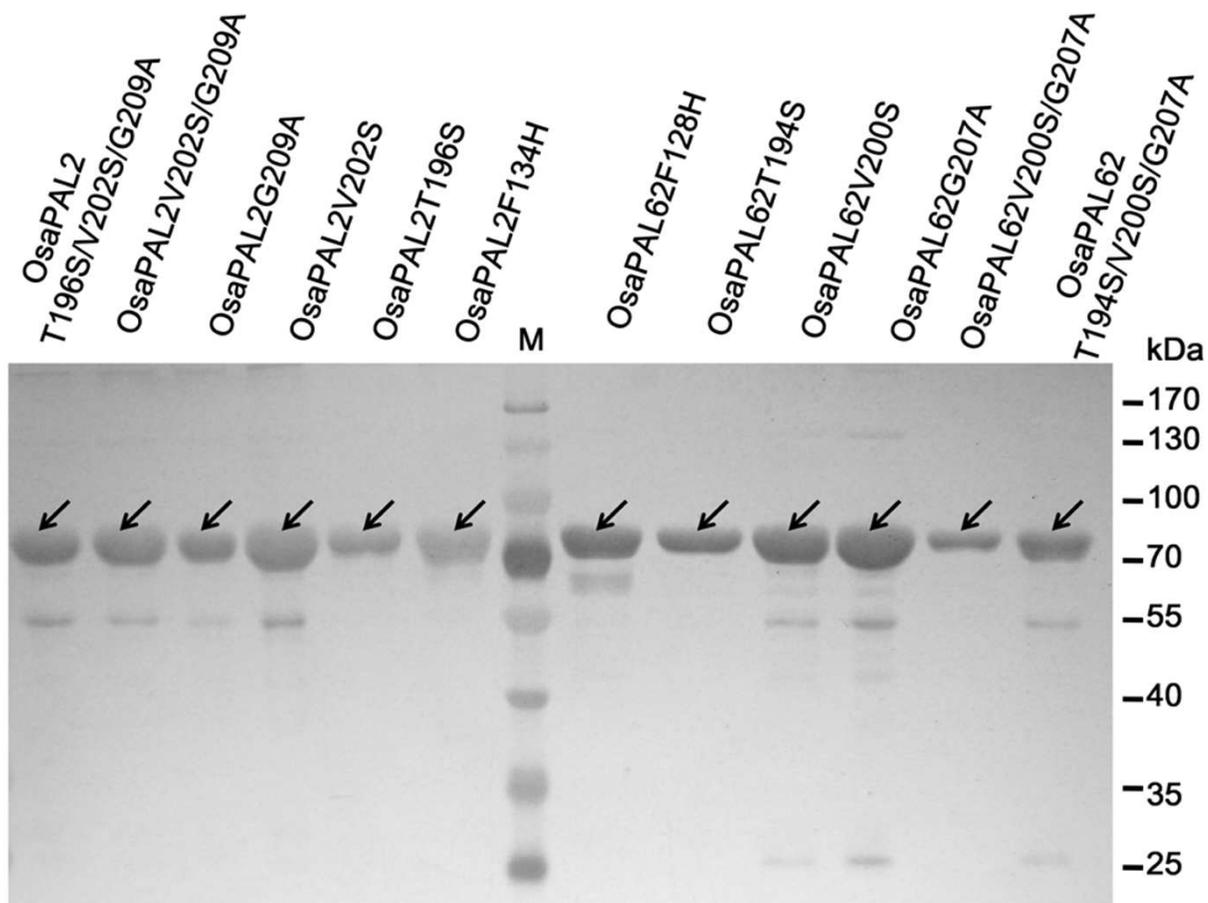
Figure 10

1
2
3



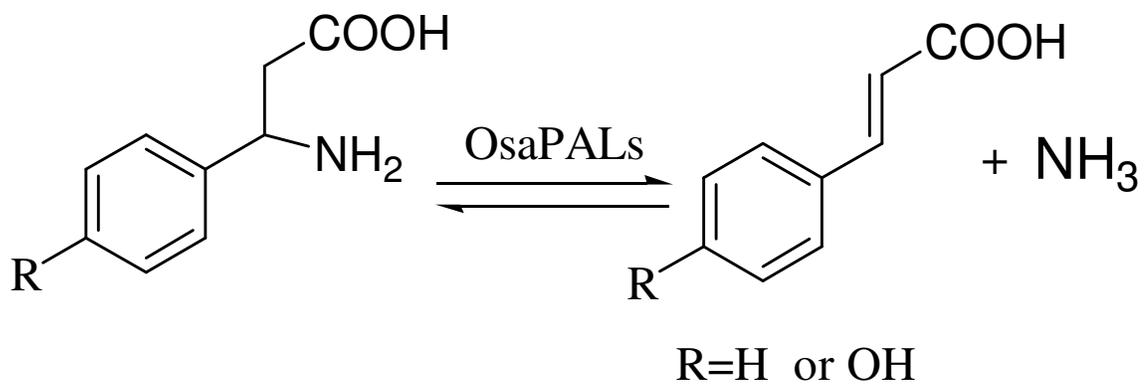
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23

Figure 11



1
2

Figure 12



Transcriptome-wide identification and characterization of *Ornithogalum saundersiae* phenylalanine ammonia-lyase gene family